

DJILLALI LIABES UNIVERSITY OF SIDI BEL ABBES  
FACULTY OF EXACT SCIENCES  
DEPARTMENT OF COMPUTER SCIENCES



*Module : Data Mining*  
1ST YEAR OF MASTER'S DEGREE IN  
NETWORKS, SYSTEMS & INFORMATION SECURITY(RSSI)  
2021/2022

---

**Data preprocessing with Weka  
TP-03**

---

*Student:*  
HADJAZI Mohammed  
Hisham  
*Group:* 01 / RSSI

*Module Instructor:*  
Pr.ELBERRICHI Zakaria  
*TP Instructor:*  
Dr.FAHSI.Mahmoud

*A paper submitted in fulfilment of the requirements for the  
Data Mining TP-03*

April 4, 2022

# Contents

<b>List of Figures</b>	<b>ix</b>
<b>1 Dataset Preparation</b>	<b>1</b>
1.1 Thyroid Disease Data Set . . . . .	1
1.1.1 Dataset Information from the owner . . . . .	1
1.2 sick dataset . . . . .	2
1.2.1 General information about the dataset . . . . .	2
<b>2 Choosing Algorithms Process</b>	<b>3</b>
2.1 Introduction . . . . .	3
2.2 One Rule "default settings" . . . . .	3
2.2.1 Cross Validation (5 Folds) . . . . .	3
2.2.2 Cross Validation (10 Folds) . . . . .	4
2.2.3 Cross Validation (15 Folds) . . . . .	5
2.2.4 Percentage Split 50% . . . . .	6
2.2.5 Percentage Split 60% . . . . .	7
2.2.6 Percentage Split 66% . . . . .	8
2.2.7 Leave one out Fold . . . . .	9
2.2.8 Conclusion . . . . .	10
2.3 Naive Bayes "default settings" . . . . .	10
2.3.1 Cross Validation (5 Folds) . . . . .	10
2.3.2 Cross Validation (10 Folds) . . . . .	13
2.3.3 Cross Validation (15 Folds) . . . . .	15
2.3.4 Percentage Split 50% . . . . .	18
2.3.5 Percentage Split 60% . . . . .	21
2.3.6 Percentage Split 66% . . . . .	23
2.3.7 Leave one out Fold . . . . .	26
2.3.8 Conclusion . . . . .	29
2.4 K-nearest Neighbour " K-value = 3 and default settings" . . . . .	29
2.4.1 Cross Validation (5 Folds) . . . . .	29
2.4.2 Cross Validation (10 Folds) . . . . .	30
2.4.3 Cross Validation (15 Folds) . . . . .	31
2.4.4 Percentage Split 50% . . . . .	32
2.4.5 Percentage Split 60% . . . . .	33
2.4.6 Percentage Split 66% . . . . .	34
2.4.7 Leave one out Fold . . . . .	34
2.4.8 Conclusion . . . . .	35
2.5 C4.5 "default settings" . . . . .	36
2.5.1 Cross Validation (5 Folds) . . . . .	36
2.5.2 Cross Validation (10 Folds) . . . . .	37
2.5.3 Cross Validation (15 Folds) . . . . .	39
2.5.4 Percentage Split 50% . . . . .	41
2.5.5 Percentage Split 60% . . . . .	42

2.5.6	Percentage Split 66% . . . . .	44
2.5.7	Leave one out Fold . . . . .	46
2.5.8	Conclusion . . . . .	47
2.6	PART "default settings" . . . . .	47
2.6.1	Cross Validation (5 Folds) . . . . .	47
2.6.2	Cross Validation (10 Folds) . . . . .	49
2.6.3	Cross Validation (15 Folds) . . . . .	51
2.6.4	Percentage Split 50% . . . . .	53
2.6.5	Percentage Split 60% . . . . .	54
2.6.6	Percentage Split 66% . . . . .	56
2.6.7	Leave one out Fold . . . . .	58
2.6.8	Conclusion . . . . .	60
2.7	Random Forest "default settings" . . . . .	60
2.7.1	Cross Validation (5 Folds) . . . . .	60
2.7.2	Cross Validation (10 Folds) . . . . .	61
2.7.3	Cross Validation (15 Folds) . . . . .	62
2.7.4	Percentage Split 50% . . . . .	63
2.7.5	Percentage Split 60% . . . . .	64
2.7.6	Percentage Split 66% . . . . .	65
2.7.7	Leave one out Fold . . . . .	65
2.7.8	Conclusion . . . . .	66
2.8	General Conclusion . . . . .	66
3	<b>Testing with preprocessed dataset and PART Algorithm</b> . . . . .	67
3.1	Non supervised discretization and Equal width binning . . . . .	67
3.1.1	2 bins . . . . .	67
	Cross Validation 5 folds . . . . .	67
	Cross Validation 10 folds . . . . .	68
	Cross Validation 15 folds . . . . .	70
	Percentage Split 50% . . . . .	71
	Percentage Split 60% . . . . .	73
	Percentage Split 66% . . . . .	74
	Leave one out fold . . . . .	76
3.1.2	5 bins . . . . .	77
	Cross Validation 5 folds . . . . .	77
	Cross Validation 10 folds . . . . .	79
	Cross Validation 15 folds . . . . .	82
	Percentage Split 50% . . . . .	84
	Percentage Split 60% . . . . .	86
	Percentage Split 66% . . . . .	88
	Leave one out fold . . . . .	90
3.1.3	10 bins . . . . .	92
	Cross Validation 5 folds . . . . .	92
	Cross Validation 10 folds . . . . .	96
	Cross Validation 15 folds . . . . .	99
	Percentage Split 50% . . . . .	102
	Percentage Split 60% . . . . .	106
	Percentage Split 66% . . . . .	109
	Leave one out fold . . . . .	112
3.1.4	40 bins . . . . .	116
	Cross Validation 5 folds . . . . .	116

Cross Validation 10 folds . . . . .	118
Cross Validation 15 folds . . . . .	120
Percentage Split 50% . . . . .	122
Percentage Split 60% . . . . .	124
Percentage Split 66% . . . . .	127
Leave one out fold . . . . .	129
3.1.5 Conclusion . . . . .	131
3.2 Non supervised discretization and Equal frequency binning . . . . .	131
3.2.1 2 bins . . . . .	131
Cross Validation 5 folds . . . . .	131
Cross Validation 10 folds . . . . .	133
Cross Validation 15 folds . . . . .	135
Percentage Split 50% . . . . .	137
Percentage Split 60% . . . . .	139
Percentage Split 66% . . . . .	141
Leave one out fold . . . . .	143
3.2.2 5 bins . . . . .	145
Cross Validation 5 folds . . . . .	145
Cross Validation 10 folds . . . . .	154
Cross Validation 15 folds . . . . .	162
Percentage Split 50% . . . . .	170
Percentage Split 60% . . . . .	179
Percentage Split 66% . . . . .	187
Leave one out fold . . . . .	196
3.2.3 10 bins . . . . .	204
Cross Validation 5 folds . . . . .	204
Cross Validation 10 folds . . . . .	208
Cross Validation 15 folds . . . . .	211
Percentage Split 50% . . . . .	215
Percentage Split 60% . . . . .	219
Percentage Split 66% . . . . .	223
Leave one out fold . . . . .	227
3.2.4 40 bins . . . . .	230
Cross Validation 5 folds . . . . .	230
Cross Validation 10 folds . . . . .	233
Cross Validation 15 folds . . . . .	235
Percentage Split 50% . . . . .	237
Percentage Split 60% . . . . .	239
Percentage Split 66% . . . . .	241
Leave one out fold . . . . .	244
3.2.5 Conclusion . . . . .	246
3.3 Supervised discretization . . . . .	246
Cross Validation 5 folds . . . . .	246
Cross Validation 10 folds . . . . .	248
Cross Validation 15 folds . . . . .	251
Percentage Split 50% . . . . .	253
Percentage Split 60% . . . . .	255
Percentage Split 66% . . . . .	257
Leave one out fold . . . . .	259
3.3.1 Conclusion . . . . .	261
3.4 Attribute Selection . . . . .	262

3.4.1	Using Filters . . . . .	262
	Cross Validation 5 folds . . . . .	264
	Cross Validation 10 folds . . . . .	265
	Cross Validation 15 folds . . . . .	266
	Percentage Split 50% . . . . .	268
	Percentage Split 60% . . . . .	269
	Percentage Split 66% . . . . .	270
	Leave one out fold . . . . .	272
3.4.2	Using Wrapper . . . . .	273
	Features selection . . . . .	274
	Cross Validation 5 folds . . . . .	275
	Cross Validation 10 folds . . . . .	276
	Cross Validation 15 folds . . . . .	278
	Percentage Split 50% . . . . .	279
	Percentage Split 60% . . . . .	280
	Percentage Split 66% . . . . .	281
	Leave one out fold . . . . .	283
3.4.3	Conclusion . . . . .	284
3.5	Data Cleaning . . . . .	284
3.5.1	Missing Values . . . . .	284
	Cross Validation 5 folds . . . . .	285
	Cross Validation 10 folds . . . . .	287
	Cross Validation 15 folds . . . . .	289
	Percentage Split 50% . . . . .	291
	Percentage Split 60% . . . . .	293
	Percentage Split 66% . . . . .	294
	Leave one out fold . . . . .	296
3.5.2	Missing Values + Extreme Values . . . . .	298
	Cross Validation 5 folds . . . . .	298
	Cross Validation 10 folds . . . . .	300
	Cross Validation 15 folds . . . . .	302
	Percentage Split 50% . . . . .	304
	Percentage Split 60% . . . . .	306
	Percentage Split 66% . . . . .	309
	Leave one out fold . . . . .	311
3.5.3	Missing Values + Outliers . . . . .	313
	Cross Validation 5 folds . . . . .	313
	Cross Validation 10 folds . . . . .	315
	Cross Validation 15 folds . . . . .	317
	Percentage Split 50% . . . . .	319
	Percentage Split 60% . . . . .	321
	Percentage Split 66% . . . . .	323
	Leave one out fold . . . . .	325
3.5.4	Conclusion . . . . .	327
3.6	Combos . . . . .	327
3.6.1	Missing Values + Outliers + Extreme Values . . . . .	327
	Cross Validation 5 folds . . . . .	327
	Cross Validation 10 folds . . . . .	329
	Cross Validation 15 folds . . . . .	330
	Percentage Split 50% . . . . .	332
	Percentage Split 60% . . . . .	334

Percentage Split 66% . . . . .	335
Leave one out fold . . . . .	337
3.6.2 Missing + Extreme + Outliers + frequency 40 . . . . .	339
Cross Validation 5 folds . . . . .	339
Cross Validation 10 folds . . . . .	341
Cross Validation 15 folds . . . . .	343
Percentage Split 50% . . . . .	345
Percentage Split 60% . . . . .	346
Percentage Split 66% . . . . .	348
Leave one out fold . . . . .	350
3.6.3 Conclusion . . . . .	352
3.7 General Conclusion . . . . .	352
<b>4 Testing with preprocessed dataset and C4.5 Algorithm . . . . .</b>	<b>354</b>
4.1 Non supervised discretization and Equal width binning . . . . .	354
4.1.1 2 bins . . . . .	354
Cross Validation 5 folds . . . . .	354
Cross Validation 10 folds . . . . .	355
Cross Validation 15 folds . . . . .	356
Percentage Split 50% . . . . .	357
Percentage Split 60% . . . . .	358
Percentage Split 66% . . . . .	359
Leave one out fold . . . . .	360
4.1.2 5 bins . . . . .	361
Cross Validation 5 folds . . . . .	361
Cross Validation 10 folds . . . . .	362
Cross Validation 15 folds . . . . .	363
Percentage Split 50% . . . . .	364
Percentage Split 60% . . . . .	365
Percentage Split 66% . . . . .	366
Leave one out fold . . . . .	367
4.1.3 10 bins . . . . .	369
Cross Validation 5 folds . . . . .	369
Cross Validation 10 folds . . . . .	371
Cross Validation 15 folds . . . . .	372
Percentage Split 50% . . . . .	374
Percentage Split 60% . . . . .	376
Percentage Split 66% . . . . .	378
Leave one out fold . . . . .	380
4.1.4 40 bins . . . . .	382
Cross Validation 5 folds . . . . .	382
Cross Validation 10 folds . . . . .	383
Cross Validation 15 folds . . . . .	384
Percentage Split 50% . . . . .	385
Percentage Split 60% . . . . .	387
Percentage Split 66% . . . . .	388
Leave one out fold . . . . .	389
4.1.5 Conclusion . . . . .	390
4.2 Non supervised discretization and Equal frequency binning . . . . .	391
4.2.1 2 bins . . . . .	391
Cross Validation 5 folds . . . . .	391

Cross Validation 10 folds . . . . .	392
Cross Validation 15 folds . . . . .	393
Percentage Split 50% . . . . .	394
Percentage Split 60% . . . . .	394
Percentage Split 66% . . . . .	395
Leave one out fold . . . . .	396
4.2.2 5 bins . . . . .	397
Cross Validation 5 folds . . . . .	397
Cross Validation 10 folds . . . . .	399
Cross Validation 15 folds . . . . .	401
Percentage Split 50% . . . . .	403
Percentage Split 60% . . . . .	405
Percentage Split 66% . . . . .	407
Leave one out fold . . . . .	409
4.2.3 10 bins . . . . .	411
Cross Validation 5 folds . . . . .	411
Cross Validation 10 folds . . . . .	412
Cross Validation 15 folds . . . . .	414
Percentage Split 50% . . . . .	416
Percentage Split 60% . . . . .	418
Percentage Split 66% . . . . .	420
Leave one out fold . . . . .	422
4.2.4 40 bins . . . . .	424
Cross Validation 5 folds . . . . .	424
Cross Validation 10 folds . . . . .	425
Cross Validation 15 folds . . . . .	426
Percentage Split 50% . . . . .	427
Percentage Split 60% . . . . .	428
Percentage Split 66% . . . . .	429
Leave one out fold . . . . .	431
4.2.5 Conclusion . . . . .	432
4.3 Supervised discretization . . . . .	432
Cross Validation 5 folds . . . . .	432
Cross Validation 10 folds . . . . .	434
Cross Validation 15 folds . . . . .	436
Percentage Split 50% . . . . .	437
Percentage Split 60% . . . . .	439
Percentage Split 66% . . . . .	441
Leave one out fold . . . . .	443
4.3.1 Conclusion . . . . .	445
4.4 Attribute Selection . . . . .	445
4.4.1 Using Filters . . . . .	445
Cross Validation 5 folds . . . . .	446
Cross Validation 10 folds . . . . .	447
Cross Validation 15 folds . . . . .	448
Percentage Split 50% . . . . .	450
Percentage Split 60% . . . . .	451
Percentage Split 66% . . . . .	452
Leave one out fold . . . . .	453
4.4.2 Using Wrapper . . . . .	454
Features selection . . . . .	455

Cross Validation 5 folds . . . . .	456
Cross Validation 10 folds . . . . .	458
Cross Validation 15 folds . . . . .	459
Percentage Split 50% . . . . .	461
Percentage Split 60% . . . . .	462
Percentage Split 66% . . . . .	463
Leave one out fold . . . . .	465
4.4.3 Conclusion . . . . .	466
4.5 Data Cleaning . . . . .	467
4.5.1 Missing Values . . . . .	467
Cross Validation 5 folds . . . . .	467
Cross Validation 10 folds . . . . .	469
Cross Validation 15 folds . . . . .	470
Percentage Split 50% . . . . .	472
Percentage Split 60% . . . . .	473
Percentage Split 66% . . . . .	475
Leave one out fold . . . . .	476
4.5.2 Extreme Values . . . . .	478
Cross Validation 5 folds . . . . .	478
Cross Validation 10 folds . . . . .	480
Cross Validation 15 folds . . . . .	481
Percentage Split 50% . . . . .	482
Percentage Split 60% . . . . .	484
Percentage Split 66% . . . . .	486
Leave one out fold . . . . .	487
4.5.3 Outliers . . . . .	489
Cross Validation 5 folds . . . . .	489
Cross Validation 10 folds . . . . .	490
Cross Validation 15 folds . . . . .	492
Percentage Split 50% . . . . .	493
Percentage Split 60% . . . . .	495
Percentage Split 66% . . . . .	496
Leave one out fold . . . . .	498
4.5.4 Conclusion . . . . .	499
4.6 Combos . . . . .	500
4.6.1 Missing Values + Outliers + Extreme Values . . . . .	500
Cross Validation 5 folds . . . . .	500
Cross Validation 10 folds . . . . .	501
Cross Validation 15 folds . . . . .	502
Percentage Split 50% . . . . .	504
Percentage Split 60% . . . . .	505
Percentage Split 66% . . . . .	507
Leave one out fold . . . . .	508
4.6.2 Missing Values + Outliers + Extreme Values + frequency 40 . . . . .	510
Cross Validation 5 folds . . . . .	510
Cross Validation 10 folds . . . . .	511
Cross Validation 15 folds . . . . .	512
Percentage Split 50% . . . . .	513
Percentage Split 60% . . . . .	514
Percentage Split 66% . . . . .	516
Leave one out fold . . . . .	517

4.6.3 Conclusion . . . . .	518
4.7 General Conclusion . . . . .	518
<b>5 Coding part . . . . .</b>	<b>520</b>
5.1 Execution with no Filters . . . . .	520
5.2 Using Discretize() Filter . . . . .	520
5.3 Using ReplaceMissingValues() Filter . . . . .	521
5.4 Using ReplaceMissingValues() and Discretize() Filter . . . . .	522
<b>6 Conclusion . . . . .</b>	<b>524</b>
<b>A Source Code . . . . .</b>	<b>526</b>
A.1 apigui.java . . . . .	526

# List of Figures

3.1	PART Filters1	262
3.2	PART Filters2	262
3.3	PART Filters3	263
3.4	PART Wrapper1	273
3.5	PART Wrapper2	274
3.6	PART Wrapper3	274
3.7	PART MissingValues	285
4.1	C4.5 Filters1	445
4.2	4.5 Filters2	446
4.3	4.5 Filters3	446
4.4	C4.5 Wrapper1	454
4.5	C4.5 Wrapper2	455
4.6	C4.5 Wrapper3	455
4.7	C4.5 MissingValues	467
5.1	No Filters	520
5.2	Discretize	521
5.3	Replace Missing Values	522
5.4	Replace Missing Values and Discretize	523

## Chapter 1

# Dataset Preparation

### 1.1 Thyroid Disease Data Set

This dataset was a part of **Thyroid Disease Databases** obtained from [openml.org](#)[9] I have chosen only 1 dataset out of the 6 that this research provided, the file that i will be using is called **sick.data.CSV** that has **3772 instances** of different patients. that are been tested on **30 different tests Attributes.**[8]

#### 1.1.1 Dataset Information from the owner

##### General Description of Thyroid Disease Databases and Related Files

This directory contains 6 databases, corresponding test set, and corresponding documentation. They were left at the University of California at Irvine by Ross Quinlan during his visit in 1987 for the 1987 Machine Learning Workshop.

The documentation files (with file extension "names") are formatted to be read by Quinlan's C4 decision tree program. Though briefer than the other documentation files found in this database repository, they should suffice to describe the database, specifically:

1. Source
2. Number and names of attributes (including class names)
3. Types of values that each attribute takes

In general, these databases are quite similar and can be characterized somewhat as follows:

1. Many attributes (29 or so, mostly the same set over all the databases)
2. mostly numeric or Boolean valued attributes
3. thyroid disease domains (records provided by the Garavan Institute of Sydney, Australia)
4. several missing attribute values (signified by "?")
5. small number of classes (under 10, changes with each database)
7. 2800 instances in each data set
8. 972 instances in each test set (It seems that the test sets' instances are disjoint with respect to the corresponding data sets, but this has not been verified)

See the following for a discussion of relevant experiments and related work:  
 Quinlan, J.R., Compton, P.J., Horn, K.A., & Lazurus, L. (1986).  
 Inductive knowledge acquisition: A case study.  
 In Proceedings of the Second Australian Conference on Applications of Expert Systems. Sydney, Australia.

Quinlan, J.R. (1986). Induction of decision trees. *Machine Learning*, 1, 81--106.

Note that the instances in these databases are followed by a vertical bar and a number. These appear to be a patient id number. The vertical bar is interepreted by Quinlan's algorithms as "ignore the remainder of this line".

---

This database now also contains an additional two data files, named **hypothyroid.data** and **sick-euthyroid.data**. They have approximately the same data format and set of attributes as the other 6 databases, but their integrity is questionable. Ross Quinlan is concerned that they may have been corrupted since they first arrived at UCI, but we have not yet established the validity of this possibility. These 2 databases differ

in terms of their number of instances (3163) and lack of corresponding test files. They each have 2 concepts (negative/hypothyroid and sick-euthyroid/negative respectively). Their source also appears to be the Garvan institute. Each contains several missing values.

Another relatively recent file thyroid0387.data has been added that contains the latest version of an archive of thyroid diagnoses obtained from the Garvan Institute, consisting of 9172 records from 1984 to early 1987.

A domain theory related to thyroid disease has also been added recently (thyroid.theory).

The files new-thyroid.[names,data] were donated by Stefan Aberhard.

## 1.2 sick dataset

### 1.2.1 General information about the dataset

		name	type	enum	ints	real	missing	distinct	(1)
1	'age'		Int	0%	100%	0%	1 / 0%	93 / 2%	0%
2	'sex'		Enum	96%	0%	0%	150 / 4%	2 / 0%	0%
3	'on thyroxine'		Enum	100%	0%	0%	0 / 0%	2 / 0%	0%
4	'query on thyroxine'		Enum	100%	0%	0%	0 / 0%	2 / 0%	0%
5	'on antithyroid medication'		Enum	100%	0%	0%	0 / 0%	2 / 0%	0%
6	'sick'		Enum	100%	0%	0%	0 / 0%	2 / 0%	0%
7	'pregnant'		Enum	100%	0%	0%	0 / 0%	2 / 0%	0%
8	'thyroid surgery'		Enum	100%	0%	0%	0 / 0%	2 / 0%	0%
9	'I131 treatment'		Enum	100%	0%	0%	0 / 0%	2 / 0%	0%
10	'query hypothyroid'		Enum	100%	0%	0%	0 / 0%	2 / 0%	0%
11	'query hyperthyroid'		Enum	100%	0%	0%	0 / 0%	2 / 0%	0%
12	'lithium'		Enum	100%	0%	0%	0 / 0%	2 / 0%	0%
13	'goitre'		Enum	100%	0%	0%	0 / 0%	2 / 0%	0%
14	'tumor'		Enum	100%	0%	0%	0 / 0%	2 / 0%	0%
15	'hypopituitary'		Enum	100%	0%	0%	0 / 0%	2 / 0%	0%
16	'psych'		Enum	100%	0%	0%	0 / 0%	2 / 0%	0%
17	'TSH measured'		Enum	100%	0%	0%	0 / 0%	2 / 0%	0%
18	'TSH'		Real	0%	11%	79%	369 / 10%	287 / 8%	2%
19	'T3 measured'		Enum	100%	0%	0%	0 / 0%	2 / 0%	0%
20	'T3'		Real	0%	9%	71%	769 / 20%	69 / 2%	0%
21	'TT4 measured'		Enum	100%	0%	0%	0 / 0%	2 / 0%	0%
22	'TT4'		Real	0%	94%	0%	231 / 6%	241 / 6%	1%
23	'T4U measured'		Enum	100%	0%	0%	0 / 0%	2 / 0%	0%
24	'T4U'		Real	0%	2%	87%	387 / 10%	146 / 4%	1%
25	'FTI measured'		Enum	100%	0%	0%	0 / 0%	2 / 0%	0%
26	'FTI'		Real	0%	90%	0%	385 / 10%	234 / 6%	2%
27	'TBG measured'		Enum	100%	0%	0%	0 / 0%	1 / 0%	0%
28	'TBG'		Real	0%	0%	100%	3772 / 100%	0 / 0%	0%
29	'referral source'		Enum	100%	0%	0%	0 / 0%	5 / 0%	0%
30	'Class'		Enum	100%	0%	0%	0 / 0%	2 / 0%	0%

## Chapter 2

# Choosing Algorithms Process

### 2.1 Introduction

Here we will conduct a series of tests to choose the best algorithm that we will use in the rest of our tests. we will be using 3 measures of evaluation with to evaluate 6 algorithms on the original dataset before any preprocessing then the two best algorithms will be used in the rest of the tests.

The results are are taken straight out of weka's classifier output and made into text files (will be included with the supplementary files). at the end of the tests there will be a conclusion of the results along with a table that summarises the performance according to correctly classified instances.

### 2.2 One Rule "default settings"

#### 2.2.1 Cross Validation (5 Folds)

```
==== Run information ====
Scheme:      weka.classifiers.rules.OneR -B 6
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   5-fold cross-validation
==== Classifier model (full training set) ====

```

```

T3:
< 0.25 -> negative
< 0.35 -> sick
< 0.55 -> negative
< 1.15 -> sick
>= 1.15 -> negative
? -> negative
(3644/3772 instances correct)

Time taken to build model: 0.17 seconds

==== Stratified cross-validation ====
==== Summary ====

  Correctly Classified Instances      3635           96.368 %
  Incorrectly Classified Instances   137            3.632 %
  Kappa statistic                   0.7083
  Mean absolute error              0.0363
  Root mean squared error          0.1906
  Relative absolute error          31.5182 %
  Root relative squared error     79.4834 %
  Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.975     0.208     0.986     0.975     0.981     0.711     0.884     0.985     negative
      0.792     0.025     0.673     0.792     0.728     0.711     0.884     0.546     sick
  Weighted Avg.       0.964     0.197     0.967     0.964     0.965     0.711     0.884     0.958

==== Confusion Matrix ====

    a     b  <- classified as
3452   89 |  a = negative
    48  183 |  b = sick

```

## 2.2.2 Cross Validation (10 Folds)

```

==== Run information ====

Scheme:      weka.classifiers.rules.OneR -B 6
Relation:    sick
Instances:   3772
Attributes:  30
             age
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             I131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary
             psych
             TSH_measured
             TSH
             T3_measured
             T3
             TT4_measured
             TT4
             T4U_measured
             T4U
             FTI_measured
             FTI
             TBG_measured
             TBG
             referral_source
             Class
Test mode:   10-fold cross-validation

==== Classifier model (full training set) ====

T3:
< 0.25 -> negative
< 0.35 -> sick
< 0.55 -> negative
< 1.15 -> sick

```

```

>= 1.15 -> negative
? -> negative
(3644/3772 instances correct)

Time taken to build model: 0.07 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3636           96.3945 %
Incorrectly Classified Instances   136            3.6055 %
Kappa statistic                   0.7066
Mean absolute error               0.0361
Root mean squared error          0.1899
Relative absolute error          31.2959 %
Root relative squared error     79.1928 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
0.976    0.221    0.985     0.976    0.981     0.709    0.878    0.984    negative
0.779    0.024    0.679     0.779    0.726     0.709    0.878    0.543    sick
Weighted Avg.                     0.964    0.209    0.967     0.964    0.965     0.709    0.878    0.957

==== Confusion Matrix ====

      a      b  <-- classified as
3456   85 |      a = negative
      51  180 |      b = sick

```

### 2.2.3 Cross Validation (15 Folds)

```

==== Run information ====

Scheme:      weka.classifiers.rules.OneR -B 6
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   15-fold cross-validation

==== Classifier model (full training set) ====

T3:
< 0.25 -> negative
< 0.35 -> sick
< 0.55 -> negative
< 1.15 -> sick
>= 1.15 -> negative
? -> negative
(3644/3772 instances correct)

```

```

Time taken to build model: 0.04 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3636           96.3945 %
Incorrectly Classified Instances   136            3.6055 %
Kappa statistic                   0.7055
Mean absolute error               0.0361
Root mean squared error          0.1899
Relative absolute error          31.298 %
Root relative squared error     79.1925 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall    F-Measure   MCC      ROC Area   PRC Area   Class
      0.976     0.225     0.985      0.976     0.981      0.707     0.876     0.984     negative
      0.775     0.024     0.681      0.775     0.725      0.707     0.876     0.541     sick
Weighted Avg.      0.964     0.213     0.967      0.964     0.965      0.707     0.876     0.957

==== Confusion Matrix ====

      a      b  <-- classified as
3457   84 |     a = negative
      52  179 |     b = sick

```

## 2.2.4 Percentage Split 50%

```

==== Run information ====

Scheme:      weka.classifiers.rules.OneR -B 6
Relation:    sick
Instances:   3772
Attributes:  30
             age
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             t131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary
             psych
             TSH_measured
             TSH
             T3_measured
             T3
             TT4_measured
             TT4
             T4U_measured
             T4U
             FTI_measured
             FTI
             TBG_measured
             TBG
             referral_source
             Class
Test mode:   split 50.0% train, remainder test

==== Classifier model (full training set) ====

T3:
< 0.25 -> negative
< 0.35 -> sick
< 0.55 -> negative
< 1.15 -> sick
>= 1.15 -> negative
? -> negative
(3644/3772 instances correct)


```

```

Time taken to build model: 0.04 seconds

==== Evaluation on test split ====
Time taken to test model on test split: 0.02 seconds

```

```

==== Summary ====
Correctly Classified Instances      1822          96.6066 %
Incorrectly Classified Instances    64           3.3934 %
Kappa statistic                   0.7238
Mean absolute error               0.0339
Root mean squared error          0.1842
Relative absolute error          29.4084 %
Root relative squared error     76.3659 %
Total Number of Instances        1886

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC   ROC Area   PRC Area   Class
0.978      0.214      0.986      0.978      0.982      0.725      0.882      0.985   negative
0.786      0.022      0.702      0.786      0.742      0.725      0.882      0.565   sick
Weighted Avg.   0.966      0.202      0.968      0.966      0.967      0.725      0.882      0.959

==== Confusion Matrix ====

      a      b  <-- classified as
1730   39 |     a = negative
      25   92 |     b = sick

```

## 2.2.5 Percentage Split 60%

```

==== Run information ====
Scheme:      weka.classifiers.rules.OneR -B 6
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 60.0% train, remainder test

==== Classifier model (full training set) ====
T3:
< 0.25 -> negative
< 0.35 -> sick
< 0.55 -> negative
< 1.15 -> sick
>= 1.15 -> negative
? -> negative
(3644/3772 instances correct)

Time taken to build model: 0.03 seconds

==== Evaluation on test split ====
Time taken to test model on test split: 0.03 seconds

==== Summary ====
Correctly Classified Instances      1458          96.6203 %

```

```

Incorrectly Classified Instances      51          3.3797 %
Kappa statistic                      0.715
Mean absolute error                  0.0338
Root mean squared error              0.1838
Relative absolute error              29.4743 %
Root relative squared error         78.0231 %
Total Number of Instances           1509

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC     ROC Area   PRC Area   Class
      0.977    0.213     0.986     0.977    0.982     0.717    0.882     0.985   negative
      0.787    0.023     0.686     0.787    0.733     0.717    0.882     0.552   sick
Weighted Avg.      0.966    0.202     0.969     0.966    0.967     0.717    0.882     0.960

==== Confusion Matrix ===

      a      b  <-- classified as
1388   32 |      a = negative
      19   70 |      b = sick

```

## 2.2.6 Percentage Split 66%

```

==== Run information ===

Scheme:      weka.classifiers.rules.OneR -B 6
Relation:    sick
Instances:   3772
Attributes:  30
             age
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             t131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary
             psych
             TSH_measured
             TSH
             T3_measured
             T3
             TT4_measured
             TT4
             T4U_measured
             T4U
             FTI_measured
             FTI
             TBG_measured
             TBG
             referral_source
             Class
Test mode:   split 66.0% train, remainder test

==== Classifier model (full training set) ===

T3:
< 0.25 -> negative
< 0.35 -> sick
< 0.55 -> negative
< 1.15 -> sick
>= 1.15 -> negative
? -> negative
(3644/3772 instances correct)

Time taken to build model: 0.04 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0.04 seconds

==== Summary ===

Correctly Classified Instances      1242      96.8799 %
Incorrectly Classified Instances     40       3.1201 %
Kappa statistic                      0.7203
Mean absolute error                  0.0312

```

```

Root mean squared error          0.1766
Relative absolute error         27.3378 %
Root relative squared error    75.2506 %
Total Number of Instances      1282

==== Detailed Accuracy By Class ====

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC     ROC Area  PRC Area  Class
      0.983    0.253    0.984    0.983    0.983    0.720    0.865    0.983    negative
      0.747    0.017    0.727    0.747    0.737    0.720    0.865    0.558    sick
Weighted Avg.      0.969    0.240    0.969    0.969    0.969    0.720    0.865    0.959

==== Confusion Matrix ====

      a      b  <-- classified as
1186  21 |      a = negative
      19  56 |      b = sick

```

### 2.2.7 Leave one out Fold

```

==== Run information ===

Scheme:      weka.classifiers.rules.OneR -B 6
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   3772-fold cross-validation

==== Classifier model (full training set) ===

T3:
< 0.25 -> negative
< 0.35 -> sick
< 0.55 -> negative
< 1.15 -> sick
>= 1.15 -> negative
? -> negative
(3644/3772 instances correct)

Time taken to build model: 0.05 seconds

==== Stratified cross-validation ===
==== Summary ===

Correctly Classified Instances      3626      96.1294 %
Incorrectly Classified Instances   146       3.8706 %
Kappa statistic                   0.6886
Mean absolute error               0.0387
Root mean squared error           0.1967
Relative absolute error           33.5948 %
Root relative squared error      82.031  %
Total Number of Instances        3772
```

```
==== Detailed Accuracy By Class ====

```

```

TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
0.974    0.229    0.985     0.974    0.979     0.691    0.872     0.984    negative
0.771    0.026    0.657     0.771    0.709     0.691    0.872     0.520    sick
Weighted Avg. 0.961    0.217    0.965     0.961    0.963     0.691    0.872     0.955

==== Confusion Matrix ====
a      b  <-- classified as
3448  93 |  a = negative
 53  178 |  b = sick

```

### 2.2.8 Conclusion

ONE RULE			
Evaluation Process		Correctly Classified Instances	notes
Cross Validation	5 folds	96.3680%	
	10 folds	96.3945%	
	15 folds	96.3945%	
Percentage Split	50.00%	96.6066%	
	60.00%	96.6203%	
	66.00%	96.8799%	BEST
Leave One Out Fold		96.1294%	

## 2.3 Naive Bayes "default settings"

### 2.3.1 Cross Validation (5 Folds)

```

==== Run information ====
Scheme:      weka.classifiers.bayes.NaiveBayes
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   5-fold cross-validation

==== Classifier model (full training set) ====
Naive Bayes Classifier

```

Attribute	Class	
	negative (0.94)	sick (0.06)
<hr/>		
age		
mean	50.99	64.0881
std. dev.	20.2327	16.3603
weight sum	3540	231
precision	4.9348	4.9348
sex		
F	2345.0	137.0
M	1054.0	90.0
[total]	3399.0	227.0
on_thyroxine		
f	3090.0	220.0
t	453.0	13.0
[total]	3543.0	233.0
query_on_thyroxine		
f	3496.0	228.0
t	47.0	5.0
[total]	3543.0	233.0
on_antithyroid_medication		
f	3499.0	232.0
t	44.0	1.0
[total]	3543.0	233.0
sick		
f	3420.0	207.0
t	123.0	26.0
[total]	3543.0	233.0
pregnant		
f	3489.0	232.0
t	54.0	1.0
[total]	3543.0	233.0
thyroid_surgery		
f	3489.0	232.0
t	54.0	1.0
[total]	3543.0	233.0
I131_treatment		
f	3484.0	231.0
t	59.0	2.0
[total]	3543.0	233.0
query_hypothyroid		
f	3336.0	204.0
t	207.0	29.0
[total]	3543.0	233.0
query_hyperthyroid		
f	3313.0	224.0
t	230.0	9.0
[total]	3543.0	233.0
lithium		
f	3525.0	231.0
t	18.0	2.0
[total]	3543.0	233.0
goitre		
f	3510.0	230.0
t	33.0	3.0
[total]	3543.0	233.0
tumor		
f	3448.0	230.0
t	95.0	3.0
[total]	3543.0	233.0
hypopituitary		
f	3542.0	231.0
t	1.0	2.0
[total]	3543.0	233.0
psych		
f	3365.0	225.0
t	178.0	8.0
[total]	3543.0	233.0
TSH_measured		
t	3175.0	230.0

```

f           368.0    3.0
[total]      3543.0   233.0

TSH
mean        5.0964  4.1999
std. dev.   25.142   13.2259
weight sum  3174     229
precision   1.8531   1.8531

T3_measured
t           2776.0   229.0
f           767.0    4.0
[total]      3543.0   233.0

T3
mean        2.108   0.8921
std. dev.   0.7813  0.4392
weight sum  2775    228
precision   0.1551  0.1551

TT4_measured
t           3311.0   232.0
f           232.0    1.0
[total]      3543.0   233.0

TT4
mean        109.4196 92.6175
std. dev.   35.6301  31.3848
weight sum  3310    231
precision   1.7833  1.7833

T4U_measured
t           3163.0   224.0
f           380.0    9.0
[total]      3543.0   233.0

T4U
mean        1.0063  0.8352
std. dev.   0.194   0.1351
weight sum  3162    223
precision   0.0143  0.0143

FTI_measured
t           3165.0   224.0
f           378.0    9.0
[total]      3543.0   233.0

FTI
mean        110.446  110.8756
std. dev.   33.0156  34.1984
weight sum  3164    223
precision   1.6867  1.6867

TBG_measured
f           3542.0   232.0
[total]      3542.0   232.0

TBG
mean        0         0
std. dev.   0.0017  0.0017
weight sum  0         0
precision   0.01     0.01

referral_source
SVHC        378.0    10.0
other       2169.0   34.0
SVI         849.0    187.0
STMW        113.0    1.0
SVHD        37.0     4.0
[total]      3546.0   236.0

Time taken to build model: 0.01 seconds

==== Stratified cross-validation ===
==== Summary ===

Correctly Classified Instances      3493          92.6034 %
Incorrectly Classified Instances   279           7.3966 %
Kappa statistic                   0.5235
Mean absolute error               0.0885
Root mean squared error           0.2289
Relative absolute error           76.788  %
Root relative squared error      95.4565 %
Total Number of Instances         3772

```

```
==== Detailed Accuracy By Class ====
      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC     ROC Area  PRC Area  Class
      0.936    0.229    0.984     0.936    0.960     0.548    0.926    0.992    negative
      0.771    0.064    0.441     0.771    0.561     0.548    0.926    0.685    sick
Weighted Avg.  0.926    0.219    0.951     0.926    0.935     0.548    0.926    0.973

==== Confusion Matrix ====
      a     b  <-- classified as
3315  226 |  a = negative
      53  178 |  b = sick
```

### 2.3.2 Cross Validation (10 Folds)

```
==== Run information ====
Scheme:      weka.classifiers.bayes.NaiveBayes
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              i131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   10-fold cross-validation

==== Classifier model (full training set) ====
Naive Bayes Classifier

          Class
Attribute    negative      sick
              (0.94)   (0.06)
=====
age
mean           50.99  64.0881
std. dev.      20.2327 16.3603
weight sum     3540    231
precision      4.9348  4.9348

sex
F              2345.0  137.0
M              1054.0   90.0
[total]        3399.0  227.0

on_thyroxine
f              3090.0  220.0
t              453.0   13.0
[total]        3543.0  233.0

query_on_thyroxine
f              3496.0  228.0
t              47.0    5.0
[total]        3543.0  233.0

on_antithyroid_medication
f              3499.0  232.0
t              44.0    1.0
```

[total]	3543.0	233.0
sick		
f	3420.0	207.0
t	123.0	26.0
[total]	3543.0	233.0
pregnant		
f	3489.0	232.0
t	54.0	1.0
[total]	3543.0	233.0
thyroid_surgery		
f	3489.0	232.0
t	54.0	1.0
[total]	3543.0	233.0
I131_treatment		
f	3484.0	231.0
t	59.0	2.0
[total]	3543.0	233.0
query_hypothyroid		
f	3336.0	204.0
t	207.0	29.0
[total]	3543.0	233.0
query_hyperthyroid		
f	3313.0	224.0
t	230.0	9.0
[total]	3543.0	233.0
lithium		
f	3525.0	231.0
t	18.0	2.0
[total]	3543.0	233.0
goitre		
f	3510.0	230.0
t	33.0	3.0
[total]	3543.0	233.0
tumor		
f	3448.0	230.0
t	95.0	3.0
[total]	3543.0	233.0
hypopituitary		
f	3542.0	231.0
t	1.0	2.0
[total]	3543.0	233.0
psych		
f	3365.0	225.0
t	178.0	8.0
[total]	3543.0	233.0
TSH_measured		
t	3175.0	230.0
f	368.0	3.0
[total]	3543.0	233.0
TSH		
mean	5.0964	4.1999
std. dev.	25.142	13.2259
weight sum	3174	229
precision	1.8531	1.8531
T3_measured		
t	2776.0	229.0
f	767.0	4.0
[total]	3543.0	233.0
T3		
mean	2.108	0.8921
std. dev.	0.7813	0.4392
weight sum	2775	228
precision	0.1551	0.1551
TT4_measured		
t	3311.0	232.0
f	232.0	1.0
[total]	3543.0	233.0
TT4		
mean	109.4196	92.6175
std. dev.	35.6301	31.3848

```

weight sum           3310      231
precision          1.7833    1.7833

T4U_measured
t                  3163.0    224.0
f                  380.0     9.0
[total]            3543.0    233.0

T4U
mean              1.0063    0.8352
std. dev.         0.194    0.1351
weight sum        3162      223
precision         0.0143    0.0143

FTI_measured
t                  3165.0    224.0
f                  378.0     9.0
[total]            3543.0    233.0

FTI
mean              110.446   110.8756
std. dev.         33.0156   34.1984
weight sum        3164      223
precision         1.6867    1.6867

TBG_measured
f                  3542.0    232.0
[total]            3542.0    232.0

TBG
mean              0          0
std. dev.         0.0017   0.0017
weight sum        0          0
precision         0.01      0.01

referral_source
SVHC              378.0     10.0
other             2169.0    34.0
SVI               849.0     187.0
STMW              113.0     1.0
SVHD              37.0      4.0
[total]            3546.0    236.0

Time taken to build model: 0.02 seconds

==== Stratified cross-validation ====
==== Summary ===

Correctly Classified Instances      3493      92.6034 %
Incorrectly Classified Instances   279       7.3966 %
Kappa statistic                   0.5249
Mean absolute error               0.0888
Root mean squared error          0.2294
Relative absolute error          77.0863 %
Root relative squared error     95.6866 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
negative  0.936    0.225    0.985    0.936    0.960    0.550    0.925    0.991    negative
sick      0.775    0.064    0.441    0.775    0.562    0.550    0.925    0.660    sick
Weighted Avg.  0.926    0.215    0.951    0.926    0.935    0.550    0.925    0.971

==== Confusion Matrix ===

      a      b  <-- classified as
3314  227 |  a = negative
      52  179 |  b = sick

```

### 2.3.3 Cross Validation (15 Folds)

```

==== Run information ===

Scheme:      weka.classifiers.bayes.NaiveBayes
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication

```

```

sick
pregnant
thyroid_surgery
I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH
T3_measured
T3
TT4_measured
TT4
T4U_measured
T4U
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Test mode: 15-fold cross-validation

==== Classifier model (full training set) ====

Naive Bayes Classifier

      Class
Attribute      negative      sick
                  (0.94)   (0.06)
-----
age
  mean           50.99  64.0881
  std. dev.     20.2327 16.3603
  weight sum    3540       231
  precision     4.9348  4.9348

sex
  F            2345.0   137.0
  M            1054.0    90.0
  [total]       3399.0   227.0

on_thyroxine
  f            3090.0   220.0
  t             453.0    13.0
  [total]       3543.0   233.0

query_on_thyroxine
  f            3496.0   228.0
  t             47.0     5.0
  [total]       3543.0   233.0

on_antithyroid_medication
  f            3499.0   232.0
  t             44.0     1.0
  [total]       3543.0   233.0

sick
  f            3420.0   207.0
  t             123.0    26.0
  [total]       3543.0   233.0

pregnant
  f            3489.0   232.0
  t              54.0     1.0
  [total]       3543.0   233.0

thyroid_surgery
  f            3489.0   232.0
  t              54.0     1.0
  [total]       3543.0   233.0

I131_treatment
  f            3484.0   231.0
  t              59.0     2.0
  [total]       3543.0   233.0

query_hypothyroid
  f            3336.0   204.0
  t              207.0    29.0
  [total]       3543.0   233.0

query_hyperthyroid

```

f	3313.0	224.0
t	230.0	9.0
[total]	3543.0	233.0
lithium		
f	3525.0	231.0
t	18.0	2.0
[total]	3543.0	233.0
goitre		
f	3510.0	230.0
t	33.0	3.0
[total]	3543.0	233.0
tumor		
f	3448.0	230.0
t	95.0	3.0
[total]	3543.0	233.0
hypopituitary		
f	3542.0	231.0
t	1.0	2.0
[total]	3543.0	233.0
psych		
f	3365.0	225.0
t	178.0	8.0
[total]	3543.0	233.0
TSH_measured		
t	3175.0	230.0
f	368.0	3.0
[total]	3543.0	233.0
TSH		
mean	5.0964	4.1999
std. dev.	25.142	13.2259
weight sum	3174	229
precision	1.8531	1.8531
T3_measured		
t	2776.0	229.0
f	767.0	4.0
[total]	3543.0	233.0
T3		
mean	2.108	0.8921
std. dev.	0.7813	0.4392
weight sum	2775	228
precision	0.1551	0.1551
TT4_measured		
t	3311.0	232.0
f	232.0	1.0
[total]	3543.0	233.0
TT4		
mean	109.4196	92.6175
std. dev.	35.6301	31.3848
weight sum	3310	231
precision	1.7833	1.7833
T4U_measured		
t	3163.0	224.0
f	380.0	9.0
[total]	3543.0	233.0
T4U		
mean	1.0063	0.8352
std. dev.	0.194	0.1351
weight sum	3162	223
precision	0.0143	0.0143
FTI_measured		
t	3165.0	224.0
f	378.0	9.0
[total]	3543.0	233.0
FTI		
mean	110.446	110.8756
std. dev.	33.0156	34.1984
weight sum	3164	223
precision	1.6867	1.6867
TBG_measured		
f	3542.0	232.0
[total]	3542.0	232.0

```

TBG
mean 0 0
std. dev. 0.0017 0.0017
weight sum 0 0
precision 0.01 0.01

referral_source
SVHC 378.0 10.0
other 2169.0 34.0
SVI 849.0 187.0
STMW 113.0 1.0
SVHD 37.0 4.0
[total] 3546.0 236.0

Time taken to build model: 0.02 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances 3501 92.8155 %
Incorrectly Classified Instances 271 7.1845 %
Kappa statistic 0.5344
Mean absolute error 0.088
Root mean squared error 0.2279
Relative absolute error 76.3613 %
Root relative squared error 95.0494 %
Total Number of Instances 3772

==== Detailed Accuracy By Class ====

      TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
0.938 0.221 0.985 0.938 0.961 0.558 0.926 0.992 negative
0.779 0.062 0.450 0.779 0.571 0.558 0.926 0.678 sick
Weighted Avg. 0.928 0.211 0.952 0.928 0.937 0.558 0.926 0.972

==== Confusion Matrix ====

     a   b  <-- classified as
3321 220 |   a = negative
      51 180 |   b = sick

```

### 2.3.4 Percentage Split 50%

```

==== Run information ====

Scheme: weka.classifiers.bayes.NaiveBayes
Relation: sick
Instances: 3772
Attributes: 30
age
sex
on_thyroxine
query_on_thyroxine
on_antithyroid_medication
sick
pregnant
thyroid_surgery
I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH
T3_measured
T3
TT4_measured
TT4
T4U_measured
T4U
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class

Test mode: split 50.0% train, remainder test

==== Classifier model (full training set) ====

```

## Naive Bayes Classifier

Attribute	Class	
	negative (0.94)	sick (0.06)
<hr/>		
age		
mean	50.99	64.0881
std. dev.	20.2327	16.3603
weight sum	3540	231
precision	4.9348	4.9348
sex		
F	2345.0	137.0
M	1054.0	90.0
[total]	3399.0	227.0
on_thyroxine		
f	3090.0	220.0
t	453.0	13.0
[total]	3543.0	233.0
query_on_thyroxine		
f	3496.0	228.0
t	47.0	5.0
[total]	3543.0	233.0
on_antithyroid_medication		
f	3499.0	232.0
t	44.0	1.0
[total]	3543.0	233.0
sick		
f	3420.0	207.0
t	123.0	26.0
[total]	3543.0	233.0
pregnant		
f	3489.0	232.0
t	54.0	1.0
[total]	3543.0	233.0
thyroid_surgery		
f	3489.0	232.0
t	54.0	1.0
[total]	3543.0	233.0
I131_treatment		
f	3484.0	231.0
t	59.0	2.0
[total]	3543.0	233.0
query_hypothyroid		
f	3336.0	204.0
t	207.0	29.0
[total]	3543.0	233.0
query_hyperthyroid		
f	3313.0	224.0
t	230.0	9.0
[total]	3543.0	233.0
lithium		
f	3525.0	231.0
t	18.0	2.0
[total]	3543.0	233.0
goitre		
f	3510.0	230.0
t	33.0	3.0
[total]	3543.0	233.0
tumor		
f	3448.0	230.0
t	95.0	3.0
[total]	3543.0	233.0
hypopituitary		
f	3542.0	231.0
t	1.0	2.0
[total]	3543.0	233.0
psych		
f	3365.0	225.0
t	178.0	8.0
[total]	3543.0	233.0

```

TSH_measured
t          3175.0    230.0
f          368.0     3.0
[total]    3543.0    233.0

TSH
mean      5.0964   4.1999
std. dev. 25.142   13.2259
weight sum 3174     229
precision 1.8531   1.8531

T3_measured
t          2776.0    229.0
f          767.0     4.0
[total]    3543.0    233.0

T3
mean      2.108    0.8921
std. dev. 0.7813   0.4392
weight sum 2775     228
precision 0.1551   0.1551

TT4_measured
t          3311.0    232.0
f          232.0     1.0
[total]    3543.0    233.0

TT4
mean      109.4196  92.6175
std. dev. 35.6301   31.3848
weight sum 3310     231
precision 1.7833   1.7833

T4U_measured
t          3163.0    224.0
f          380.0     9.0
[total]    3543.0    233.0

T4U
mean      1.0063   0.8352
std. dev. 0.194   0.1351
weight sum 3162     223
precision 0.0143   0.0143

FTI_measured
t          3165.0    224.0
f          378.0     9.0
[total]    3543.0    233.0

FTI
mean      110.446   110.8756
std. dev. 33.0156   34.1984
weight sum 3164     223
precision 1.6867   1.6867

TBG_measured
f          3542.0    232.0
[total]    3542.0    232.0

TBG
mean      0         0
std. dev. 0.0017   0.0017
weight sum 0         0
precision 0.01     0.01

referral_source
SVHC       378.0    10.0
other      2169.0   34.0
SVI        849.0    187.0
STMW       113.0     1.0
SVHD       37.0     4.0
[total]    3546.0   236.0

Time taken to build model: 0.01 seconds
==== Evaluation on test split ====
Time taken to test model on test split: 0.02 seconds
==== Summary ====
Correctly Classified Instances      1778      94.2736 %
Incorrectly Classified Instances    108       5.7264 %
Kappa statistic                      0.5952

```

```

Mean absolute error          0.0754
Root mean squared error     0.2065
Relative absolute error     65.3258 %
Root relative squared error 85.5901 %
Total Number of Instances   1886

== Detailed Accuracy By Class ==

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
0.954      0.231    0.984     0.954    0.969     0.608    0.939     0.993    negative
0.769      0.046    0.526     0.769    0.625     0.608    0.939     0.730    sick
Weighted Avg.   0.943    0.219    0.956     0.943    0.948     0.608    0.939     0.977

== Confusion Matrix ==

  a   b  <-- classified as
1688 81 |   a = negative
 27 90 |   b = sick

```

### 2.3.5 Percentage Split 60%

```

== Run information ==

Scheme:      weka.classifiers.bayes.NaiveBayes
Relation:    sick
Instances:   3772
Attributes:  30
             age
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             i131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary
             psych
             TSH_measured
             TSH
             T3_measured
             T3
             TT4_measured
             TT4
             T4U_measured
             T4U
             FTI_measured
             FTI
             TBG_measured
             TBG
             referral_source
             Class
Test mode:   split 60.0% train, remainder test

== Classifier model (full training set) ==

Naive Bayes Classifier

      Class
Attribute      negative      sick
              (0.94)   (0.06)
=====
age
mean           50.99  64.0881
std. dev.      20.2327 16.3603
weight sum     3540    231
precision      4.9348  4.9348

sex
F              2345.0   137.0
M              1054.0   90.0
[total]        3399.0   227.0

on_thyroxine
f              3090.0   220.0
t              453.0    13.0
[total]        3543.0   233.0

query_on_thyroxine
f              3496.0   228.0

```

t	47.0	5.0
[total]	3543.0	233.0
on_antithyroid_medication		
f	3499.0	232.0
t	44.0	1.0
[total]	3543.0	233.0
sick		
f	3420.0	207.0
t	123.0	26.0
[total]	3543.0	233.0
pregnant		
f	3489.0	232.0
t	54.0	1.0
[total]	3543.0	233.0
thyroid_surgery		
f	3489.0	232.0
t	54.0	1.0
[total]	3543.0	233.0
I131_treatment		
f	3484.0	231.0
t	59.0	2.0
[total]	3543.0	233.0
query_hypothyroid		
f	3336.0	204.0
t	207.0	29.0
[total]	3543.0	233.0
query_hyperthyroid		
f	3313.0	224.0
t	230.0	9.0
[total]	3543.0	233.0
lithium		
f	3525.0	231.0
t	18.0	2.0
[total]	3543.0	233.0
goitre		
f	3510.0	230.0
t	33.0	3.0
[total]	3543.0	233.0
tumor		
f	3448.0	230.0
t	95.0	3.0
[total]	3543.0	233.0
hypopituitary		
f	3542.0	231.0
t	1.0	2.0
[total]	3543.0	233.0
psych		
f	3365.0	225.0
t	178.0	8.0
[total]	3543.0	233.0
TSH_measured		
t	3175.0	230.0
f	368.0	3.0
[total]	3543.0	233.0
TSH		
mean	5.0964	4.1999
std. dev.	25.142	13.2259
weight sum	3174	229
precision	1.8531	1.8531
T3_measured		
t	2776.0	229.0
f	767.0	4.0
[total]	3543.0	233.0
T3		
mean	2.108	0.8921
std. dev.	0.7813	0.4392
weight sum	2775	228
precision	0.1551	0.1551
TT4_measured		
t	3311.0	232.0

```

f                      232.0      1.0
[total]                3543.0    233.0

TT4
mean                  109.4196  92.6175
std. dev.              35.6301   31.3848
weight sum              3310       231
precision               1.7833   1.7833

T4U_measured
t                      3163.0    224.0
f                      380.0     9.0
[total]                3543.0    233.0

T4U
mean                  1.0063    0.8352
std. dev.              0.194     0.1351
weight sum              3162      223
precision               0.0143   0.0143

FTI_measured
t                      3165.0    224.0
f                      378.0     9.0
[total]                3543.0    233.0

FTI
mean                  110.446   110.8756
std. dev.              33.0156   34.1984
weight sum              3164      223
precision               1.6867   1.6867

TBG_measured
f                      3542.0    232.0
[total]                3542.0    232.0

TBG
mean                  0          0
std. dev.              0.0017   0.0017
weight sum              0          0
precision               0.01      0.01

referral_source
SVHC                  378.0    10.0
other                 2169.0   34.0
SVI                   849.0    187.0
STMW                  113.0     1.0
SVHD                  37.0     4.0
[total]                3546.0   236.0

Time taken to build model: 0.03 seconds

==== Evaluation on test split ====

Time taken to test model on test split: 0.03 seconds

==== Summary ====

Correctly Classified Instances      1410      93.4394 %
Incorrectly Classified Instances    99       6.5606 %
Kappa statistic                   0.5379
Mean absolute error               0.0846
Root mean squared error           0.2239
Relative absolute error            73.8139 %
Root relative squared error       95.0267 %
Total Number of Instances         1509

==== Detailed Accuracy By Class ====

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
0      0.946   0.258    0.983    0.946    0.964    0.555    0.917    0.989  negative
1      0.742   0.054    0.465    0.742    0.571    0.555    0.917    0.691   sick
Weighted Avg.      0.934   0.246    0.953    0.934    0.941    0.555    0.917    0.971

==== Confusion Matrix ====

      a      b  <-- classified as
1344  76 |      a = negative
      23  66 |      b = sick

```

### 2.3.6 Percentage Split 66%

```
==== Run information ====
```

```

Scheme:      weka.classifiers.bayes.NaiveBayes
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 66.0% train, remainder test

==== Classifier model (full training set) ===

Naive Bayes Classifier

          Class
Attribute      negative      sick
                  (0.94)     (0.06)
=====
age
mean           50.99    64.0881
std. dev.      20.2327   16.3603
weight sum     3540      231
precision      4.9348   4.9348

sex
F              2345.0   137.0
M              1054.0   90.0
[total]        3399.0   227.0

on_thyroxine
f              3090.0   220.0
t              453.0    13.0
[total]        3543.0   233.0

query_on_thyroxine
f              3496.0   228.0
t              47.0     5.0
[total]        3543.0   233.0

on_antithyroid_medication
f              3499.0   232.0
t              44.0     1.0
[total]        3543.0   233.0

sick
f              3420.0   207.0
t              123.0    26.0
[total]        3543.0   233.0

pregnant
f              3489.0   232.0
t              54.0     1.0
[total]        3543.0   233.0

thyroid_surgery
f              3489.0   232.0
t              54.0     1.0
[total]        3543.0   233.0

I131_treatment
f              3484.0   231.0

```

t	59.0	2.0
[total]	3543.0	233.0
query_hypothyroid		
f	3336.0	204.0
t	207.0	29.0
[total]	3543.0	233.0
query_hyperthyroid		
f	3313.0	224.0
t	230.0	9.0
[total]	3543.0	233.0
lithium		
f	3525.0	231.0
t	18.0	2.0
[total]	3543.0	233.0
goitre		
f	3510.0	230.0
t	33.0	3.0
[total]	3543.0	233.0
tumor		
f	3448.0	230.0
t	95.0	3.0
[total]	3543.0	233.0
hypopituitary		
f	3542.0	231.0
t	1.0	2.0
[total]	3543.0	233.0
psych		
f	3365.0	225.0
t	178.0	8.0
[total]	3543.0	233.0
TSH_measured		
t	3175.0	230.0
f	368.0	3.0
[total]	3543.0	233.0
TSH		
mean	5.0964	4.1999
std. dev.	25.142	13.2259
weight sum	3174	229
precision	1.8531	1.8531
T3_measured		
t	2776.0	229.0
f	767.0	4.0
[total]	3543.0	233.0
T3		
mean	2.108	0.8921
std. dev.	0.7813	0.4392
weight sum	2775	228
precision	0.1551	0.1551
TT4_measured		
t	3311.0	232.0
f	232.0	1.0
[total]	3543.0	233.0
TT4		
mean	109.4196	92.6175
std. dev.	35.6301	31.3848
weight sum	3310	231
precision	1.7833	1.7833
T4U_measured		
t	3163.0	224.0
f	380.0	9.0
[total]	3543.0	233.0
T4U		
mean	1.0063	0.8352
std. dev.	0.194	0.1351
weight sum	3162	223
precision	0.0143	0.0143
FTI_measured		
t	3165.0	224.0
f	378.0	9.0
[total]	3543.0	233.0

```

FTI
mean 110.446 110.8756
std. dev. 33.0156 34.1984
weight sum 3164 223
precision 1.6867 1.6867

TBG_measured
f 3542.0 232.0
[total] 3542.0 232.0

TBG
mean 0 0
std. dev. 0.0017 0.0017
weight sum 0 0
precision 0.01 0.01

referral_source
SVHC 378.0 10.0
other 2169.0 34.0
SVI 849.0 187.0
STMW 113.0 1.0
SVHD 37.0 4.0
[total] 3546.0 236.0

Time taken to build model: 0.01 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0.03 seconds

==== Summary ===

Correctly Classified Instances 1194 93.1357 %
Incorrectly Classified Instances 88 6.8643 %
Kappa statistic 0.5385
Mean absolute error 0.088
Root mean squared error 0.2271
Relative absolute error 77.1257 %
Root relative squared error 96.7281 %
Total Number of Instances 1282

==== Detailed Accuracy By Class ===

      TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
0 0.940 0.213 0.986 0.940 0.963 0.563 0.920 0.988 negative
1 0.787 0.060 0.450 0.787 0.573 0.563 0.920 0.691 sick
Weighted Avg. 0.931 0.204 0.955 0.931 0.940 0.563 0.920 0.971

==== Confusion Matrix ===

     a   b  <-- classified as
1135 72 |   a = negative
    16 59 |   b = sick

```

### 2.3.7 Leave one out Fold

```

==== Run information ===

Scheme: weka.classifiers.bayes.NaiveBayes
Relation: sick
Instances: 3772
Attributes: 30
age
sex
on_thyroxine
query_on_thyroxine
on_antithyroid_medication
sick
pregnant
thyroid_surgery
I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH
T3_measured
T3

```

```

TT4_measured
TT4
T4U_measured
T4U
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Test mode: 3772-fold cross-validation

==== Classifier model (full training set) ===

Naive Bayes Classifier

          Class
Attribute      negative      sick
                  (0.94)   (0.06)
=====
age
mean           50.99  64.0881
std. dev.     20.2327 16.3603
weight sum     3540    231
precision      4.9348  4.9348

sex
F              2345.0  137.0
M              1054.0  90.0
[total]        3399.0  227.0

on_thyroxine
f              3090.0  220.0
t              453.0   13.0
[total]        3543.0  233.0

query_on_thyroxine
f              3496.0  228.0
t              47.0    5.0
[total]        3543.0  233.0

on_antithyroid_medication
f              3499.0  232.0
t              44.0    1.0
[total]        3543.0  233.0

sick
f              3420.0  207.0
t              123.0   26.0
[total]        3543.0  233.0

pregnant
f              3489.0  232.0
t              54.0    1.0
[total]        3543.0  233.0

thyroid_surgery
f              3489.0  232.0
t              54.0    1.0
[total]        3543.0  233.0

I131_treatment
f              3484.0  231.0
t              59.0    2.0
[total]        3543.0  233.0

query_hypothyroid
f              3336.0  204.0
t              207.0   29.0
[total]        3543.0  233.0

query_hyperthyroid
f              3313.0  224.0
t              230.0   9.0
[total]        3543.0  233.0

lithium
f              3525.0  231.0
t              18.0    2.0
[total]        3543.0  233.0

goitre
f              3510.0  230.0
t              33.0    3.0
[total]        3543.0  233.0

tumor

```

f	3448.0	230.0
t	95.0	3.0
[total]	3543.0	233.0
 hypopituitary		
f	3542.0	231.0
t	1.0	2.0
[total]	3543.0	233.0
 psych		
f	3365.0	225.0
t	178.0	8.0
[total]	3543.0	233.0
 TSH_measured		
t	3175.0	230.0
f	368.0	3.0
[total]	3543.0	233.0
 TSH		
mean	5.0964	4.1999
std. dev.	25.142	13.2259
weight sum	3174	229
precision	1.8531	1.8531
 T3_measured		
t	2776.0	229.0
f	767.0	4.0
[total]	3543.0	233.0
 T3		
mean	2.108	0.8921
std. dev.	0.7813	0.4392
weight sum	2775	228
precision	0.1551	0.1551
 TT4_measured		
t	3311.0	232.0
f	232.0	1.0
[total]	3543.0	233.0
 TT4		
mean	109.4196	92.6175
std. dev.	35.6301	31.3848
weight sum	3310	231
precision	1.7833	1.7833
 T4U_measured		
t	3163.0	224.0
f	380.0	9.0
[total]	3543.0	233.0
 T4U		
mean	1.0063	0.8352
std. dev.	0.194	0.1351
weight sum	3162	223
precision	0.0143	0.0143
 FTI_measured		
t	3165.0	224.0
f	378.0	9.0
[total]	3543.0	233.0
 FTI		
mean	110.446	110.8756
std. dev.	33.0156	34.1984
weight sum	3164	223
precision	1.6867	1.6867
 TBG_measured		
f	3542.0	232.0
[total]	3542.0	232.0
 TBG		
mean	0	0
std. dev.	0.0017	0.0017
weight sum	0	0
precision	0.01	0.01
 referral_source		
SVHC	378.0	10.0
other	2169.0	34.0
SVI	849.0	187.0
STMW	113.0	1.0
SVHD	37.0	4.0
[total]	3546.0	236.0

```

Time taken to build model: 0.01 seconds

==== Stratified cross-validation ====
==== Summary ===

    Correctly Classified Instances      3503           92.8685 %
    Incorrectly Classified Instances   269            7.1315 %
    Kappa statistic                   0.5364
    Mean absolute error              0.0879
    Root mean squared error          0.2288
    Relative absolute error          76.2846 %
    Root relative squared error     95.3892 %
    Total Number of Instances       3772

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC     ROC Area   PRC Area   Class
      0.938     0.221     0.985      0.938     0.961     0.560    0.925      0.992    negative
      0.779     0.062     0.452      0.779     0.572     0.560    0.925      0.656    sick
Weighted Avg.   0.929     0.211     0.952      0.929     0.937     0.560    0.925      0.971

==== Confusion Matrix ===

      a     b  <-- classified as
3323  218 |     a = negative
      51  180 |     b = sick

```

### 2.3.8 Conclusion

NAIVE BAYES			
Evaluation Process		Correctly Classified Instances	notes
Cross Validation	5 folds	92.6034%	
	10 folds	92.6034%	
	15 folds	92.8155%	
Percentage Split	50.00%	94.2736%	BEST
	60.00%	93.4394%	
	66.00%	93.1357%	
Leave One Out Fold		92.8685%	

## 2.4 K-nearest Neighbour "K-value = 3 and default settings"

### 2.4.1 Cross Validation (5 Folds)

```

==== Run information ===

Scheme:      weka.classifiers.lazy.IBk -K 3 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A \"weka.core.EuclideanDistance -R first-last\"
Relation:    sick
Instances:   3772
Attributes:  30
             age
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             I131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary
             psych
             TSH_measured
             TSH
             T3_measured
             T3

```

```

TT4_measured
TT4
T4U_measured
T4U
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Test mode: 5-fold cross-validation

==== Classifier model (full training set) ===

IB1 instance-based classifier
using 3 nearest neighbour(s) for classification

Time taken to build model: 0 seconds

==== Stratified cross-validation ===
==== Summary ===

Correctly Classified Instances      3629          96.2089 %
Incorrectly Classified Instances   143           3.7911 %
Kappa statistic                   0.6186
Mean absolute error               0.0471
Root mean squared error          0.1752
Relative absolute error          40.8833 %
Root relative squared error     73.0645 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall    F-Measure   MCC     ROC Area   PRC Area   Class
      0.989     0.455     0.971     0.989     0.980     0.629     0.877     0.985   negative
      0.545     0.011     0.768     0.545     0.638     0.629     0.877     0.628   sick
Weighted Avg.       0.962     0.427     0.958     0.962     0.959     0.629     0.877     0.963

==== Confusion Matrix ===

      a     b  <-- classified as
3503  38 |   a = negative
 105 126 |   b = sick

```

## 2.4.2 Cross Validation (10 Folds)

```

==== Run information ===

Scheme:      weka.classifiers.lazy.IBk -K 3 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A \"weka.core.EuclideanDistance -R first-last\"
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode: 10-fold cross-validation

```

```

==== Classifier model (full training set) ====
IB1 instance-based classifier
using 3 nearest neighbour(s) for classification

Time taken to build model: 0.02 seconds

==== Stratified cross-validation ====
==== Summary ===

Correctly Classified Instances      3628          96.1824 %
Incorrectly Classified Instances   144           3.8176 %
Kappa statistic                   0.6168
Mean absolute error               0.0455
Root mean squared error          0.1745
Relative absolute error           39.5034 %
Root relative squared error     72.7836 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall    F-Measure   MCC    ROC Area   PRC Area   Class
      0.989     0.455     0.971     0.989     0.980     0.627     0.874     0.984     negative
      0.545     0.011     0.764     0.545     0.636     0.627     0.874     0.617     sick
Weighted Avg.      0.962     0.427     0.958     0.962     0.959     0.627     0.874     0.962

==== Confusion Matrix ====

      a      b  <-- classified as
3502  39 |    a = negative
 105 126 |    b = sick

```

### 2.4.3 Cross Validation (15 Folds)

```

==== Run information ====
Scheme:      weka.classifiers.lazy.IBk -K 3 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A \"weka.core.EuclideanDistance -R first-last\"
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   15-fold cross-validation

==== Classifier model (full training set) ====
IB1 instance-based classifier
using 3 nearest neighbour(s) for classification

Time taken to build model: 0 seconds

==== Stratified cross-validation ====
==== Summary ===

```

```

Correctly Classified Instances      3623          96.0498 %
Incorrectly Classified Instances   149           3.9502 %
Kappa statistic                   0.6099
Mean absolute error               0.046
Root mean squared error          0.1775
Relative absolute error           39.9635 %
Root relative squared error      74.0394 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC     ROC Area   PRC Area   Class
      0.987     0.450     0.971      0.987     0.979     0.617    0.860      0.983    negative
      0.550     0.013     0.738      0.550     0.630     0.617    0.860      0.617    sick
Weighted Avg.      0.960     0.423     0.957      0.960     0.958     0.617    0.860      0.960

==== Confusion Matrix ====

      a      b  <-- classified as
3496   45 |      a = negative
 104   127 |      b = sick

```

#### 2.4.4 Percentage Split 50%

```

==== Run information ===

Scheme:      weka.classifiers.lazy.IBk -K 3 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A \"weka.core.EuclideanDistance -R first-last\"
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 50.0% train, remainder test

==== Classifier model (full training set) ===

IB1 instance-based classifier
using 3 nearest neighbour(s) for classification

Time taken to build model: 0 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 1.14 seconds

==== Summary ===

Correctly Classified Instances      1803          95.5992 %
Incorrectly Classified Instances   83           4.4008 %
Kappa statistic                   0.4838
Mean absolute error               0.0565
Root mean squared error          0.1919
Relative absolute error           48.9959 %
Root relative squared error      79.5412 %
Total Number of Instances        1886

```

```
==== Detailed Accuracy By Class ====
      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC    ROC Area  PRC Area  Class
      0.995   0.641    0.959     0.995   0.977    0.532   0.833    0.979    negative
      0.359   0.005    0.840     0.359   0.503    0.532   0.833    0.497    sick
Weighted Avg.  0.956   0.602    0.952     0.956   0.948    0.532   0.833    0.949

==== Confusion Matrix ====
      a      b  <-- classified as
1761    8 |  a = negative
    75   42 |  b = sick
```

## 2.4.5 Percentage Split 60%

```
==== Run information ====
Scheme:      weka.classifiers.lazy.IBk -K 3 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A \"weka.core.EuclideanDistance -R first-last\"
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              t131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 60.0% train, remainder test

==== Classifier model (full training set) ====
IB1 instance-based classifier
using 3 nearest neighbour(s) for classification

Time taken to build model: 0 seconds

==== Evaluation on test split ====
Time taken to test model on test split: 1.11 seconds

==== Summary ====
Correctly Classified Instances      1446          95.825 %
Incorrectly Classified Instances   63           4.175 %
Kappa statistic                   0.5002
Mean absolute error               0.0527
Root mean squared error          0.1872
Relative absolute error          45.9627 %
Root relative squared error     79.449 %
Total Number of Instances        1509

==== Detailed Accuracy By Class ====
      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC    ROC Area  PRC Area  Class
      0.994   0.618    0.963     0.994   0.978    0.539   0.829    0.979    negative
      0.382   0.006    0.810     0.382   0.519    0.539   0.829    0.501    sick
Weighted Avg.  0.958   0.582    0.953     0.958   0.951    0.539   0.829    0.951

==== Confusion Matrix ====
      a      b  <-- classified as
1446    63 |  a = negative
    63   1509 |  b = sick
```

```
a      b    <-- classified as
1412   8 |     a = negative
 55   34 |     b = sick
```

## 2.4.6 Percentage Split 66%

```
==== Run information ===

Scheme:      weka.classifiers.lazy.IBk -K 3 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A \"weka.core.EuclideanDistance -R first-last\"
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 66.0% train, remainder test

==== Classifier model (full training set) ===

IB1 instance-based classifier
using 3 nearest neighbour(s) for classification

Time taken to build model: 0 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 1.04 seconds

==== Summary ===

  Correctly Classified Instances      1229           95.8658 %
  Incorrectly Classified Instances    53            4.1342 %
  Kappa statistic                   0.5118
  Mean absolute error               0.0516
  Root mean squared error          0.1852
  Relative absolute error          45.2113 %
  Root relative squared error      78.9182 %
  Total Number of Instances        1282

==== Detailed Accuracy By Class ===

          TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
          0.993    0.600    0.964     0.993    0.978     0.544    0.834     0.980    negative
          0.400    0.007    0.789     0.400    0.531     0.544    0.834     0.512    sick
Weighted Avg.  0.959    0.565    0.954     0.959    0.952     0.544    0.834     0.953

==== Confusion Matrix ===

a      b    <-- classified as
1199   8 |     a = negative
 45   30 |     b = sick
```

## 2.4.7 Leave one out Fold

```
==== Run information ===
```

```

Scheme:      weka.classifiers.lazy.IBk -K 3 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A \\"weka.core.EuclideanDistance -R first-last\\"
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              t131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   3772-fold cross-validation

==== Classifier model (full training set) ===

IB1 instance-based classifier
using 3 nearest neighbour(s) for classification

Time taken to build model: 0 seconds

==== Stratified cross-validation ===
==== Summary ===

Correctly Classified Instances      3628          96.1824 %
Incorrectly Classified Instances   144           3.8176 %
Kappa statistic                   0.6221
Mean absolute error               0.0457
Root mean squared error          0.1767
Relative absolute error           39.6464 %
Root relative squared error      73.6737 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC     ROC Area  PRC Area  Class
      0.988    0.442    0.972     0.988    0.980     0.630    0.869     0.983    negative
      0.558    0.012    0.754     0.558    0.642     0.630    0.869     0.594    sick
Weighted Avg.      0.962    0.415    0.958     0.962    0.959     0.630    0.869     0.960

==== Confusion Matrix ===

      a      b  <-- classified as
3499   42 |    a = negative
102    129 |    b = sick

```

## 2.4.8 Conclusion

KNN			
Evaluation Process		Correctly Classified Instances	notes
Cross Validation	5 folds	96.2089%	BEST
	10 folds	96.1824%	
	15 folds	96.0498%	
Percentage Split	50.00%	95.5992%	
	60.00%	95.8250%	
	66.00%	95.8658%	
Leave One Out Fold		96.1824%	

## 2.5 C4.5 "default settings"

### 2.5.1 Cross Validation (5 Folds)

```

==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class

Test mode:   5-fold cross-validation

==== Classifier model (full training set) ===

J48 pruned tree
-----

T3 <= 1.1
|   T3_measured = t
|   |   TT4 <= 53
|   |   |   TSH <= 19
|   |   |   |   FTI <= 78
|   |   |   |   |   T4U <= 0.67
|   |   |   |   |   |   T3 <= 0.5: sick (3.0/1.0)
|   |   |   |   |   |   T3 > 0.5: negative (2.0)
|   |   |   |   |   |   T4U > 0.67: sick (7.0)
|   |   |   |   |   |   FTI > 78: negative (8.0)
|   |   |   |   |   |   TSH > 19: negative (53.0)
|   |   TT4 > 53
|   |   |   TT4 <= 136
|   |   |   |   referral_source = SVHC: sick (9.0)
|   |   |   |   referral_source = other
|   |   |   |   |   sick = f
|   |   |   |   |   |   age <= 77
|   |   |   |   |   |   |   query_hypothyroid = f
|   |   |   |   |   |   |   on_thyroxine = f
|   |   |   |   |   |   |   |   T3 <= 0.5: sick (5.0/1.0)

```

```

|   |   |   |   |   |   |   | T3 > 0.5: negative (21.0/1.0)
|   |   |   |   |   |   | on_thyroxine = t: sick (5.0/2.0)
|   |   |   |   |   |   | query_hypothyroid = t: sick (4.0/1.0)
|   |   |   |   |   |   | age > 77: sick (8.0)
|   |   |   |   |   |   | sick = t: sick (3.0)
|   |   |   |   |   |   | referral_source = SVI
|   |   |   |   |   |   | FTI <= 151
|   |   |   |   |   |   | FTI <= 69
|   |   |   |   |   |   | FTI <= 61: negative (2.08/0.08)
|   |   |   |   |   |   | FTI > 61: sick (6.22)
|   |   |   |   |   |   | FTI > 69: sick (148.36)
|   |   |   |   |   |   | FTI > 151
|   |   |   |   |   |   | age <= 73: negative (5.28/1.28)
|   |   |   |   |   |   | age > 73: sick (4.06)
|   |   |   |   |   |   | referral_source = STMW: sick (0.0)
|   |   |   |   |   |   | referral_source = SVHD: sick (3.0)
|   |   |   |   TT4 > 136
|   |   |   |   |   T4U_measured = t: negative (12.0/1.0)
|   |   |   |   |   T4U_measured = f: sick (2.0)
|   |   |   |   T3_measured = f: negative (79.64/0.31)

T3 > 1.1
|   FTI <= 171
|   |   TT4 <= 56
|   |   |   referral_source = SVHC: negative (3.89)
|   |   |   referral_source = other: negative (32.25/0.9)
|   |   |   referral_source = SVI
|   |   |   |   TSH <= 14: sick (7.79/1.0)
|   |   |   |   TSH > 14: negative (4.98)
|   |   |   |   referral_source = STMW: negative (2.03)
|   |   |   |   referral_source = SVHD: negative (0.01)
|   |   |   TT4 > 56: negative (3187.63/1.0)
|   FTI > 171
|   |   T3 <= 2.6
|   |   |   on_thyroxine = f
|   |   |   |   T3_measured = t
|   |   |   |   |   TSH <= 0.27: sick (19.02/3.02)
|   |   |   |   |   TSH > 0.27: negative (11.76)
|   |   |   |   |   T3_measured = f: negative (8.79)
|   |   |   |   |   on_thyroxine = t: negative (19.64)
|   |   |   T3 > 2.6: negative (83.58)

Number of Leaves : 34
Size of the tree : 61

Time taken to build model: 0.18 seconds

==== Stratified cross-validation ====
==== Summary ====

  Correctly Classified Instances      3713          98.4358 %
  Incorrectly Classified Instances     59           1.5642 %
  Kappa statistic                      0.8637
  Mean absolute error                  0.0182
  Root mean squared error              0.1178
  Relative absolute error             15.7674 %
  Root relative squared error         49.1385 %
  Total Number of Instances            3772

==== Detailed Accuracy By Class ====

          TP Rate  FP Rate  Precision  Recall   F-Measure  MCC    ROC Area  PRC Area  Class
          0.992    0.130    0.992    0.992    0.992    0.864    0.949    0.992    negative
          0.870    0.008    0.874    0.870    0.872    0.864    0.950    0.812    sick
  Weighted Avg.    0.984    0.122    0.984    0.984    0.984    0.864    0.949    0.981

==== Confusion Matrix ====

      a     b  <- classified as
3512  29 |     a = negative
      30  201 |     b = sick

```

## 2.5.2 Cross Validation (10 Folds)

```

==== Run information ====

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine

```

```

on_antithyroid_medication
sick
pregnant
thyroid_surgery
I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH
T3_measured
T3
TT4_measured
TT4
T4U_measured
T4U
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Test mode: 10-fold cross-validation

==== Classifier model (full training set) ====

J48 pruned tree
-----
T3 <= 1.1
| T3_measured = t
| | TT4 <= 53
| | | TSH <= 19
| | | | FTI <= 78
| | | | | T4U <= 0.67
| | | | | | T3 <= 0.5: sick (3.0/1.0)
| | | | | | T3 > 0.5: negative (2.0)
| | | | | | T4U > 0.67: sick (7.0)
| | | | | | FTI > 78: negative (8.0)
| | | | | | TSH > 19: negative (53.0)
| | TT4 > 53
| | | TT4 <= 136
| | | | referral_source = SVHC: sick (9.0)
| | | | referral_source = other
| | | | | sick = f
| | | | | | age <= 77
| | | | | | | query_hypothyroid = f
| | | | | | | | on_thyroxine = f
| | | | | | | | | T3 <= 0.5: sick (5.0/1.0)
| | | | | | | | | T3 > 0.5: negative (21.0/1.0)
| | | | | | | | | on_thyroxine = t: sick (5.0/2.0)
| | | | | | | | | query_hypothyroid = t: sick (4.0/1.0)
| | | | | | | | | age > 77: sick (8.0)
| | | | | | | | | sick = t: sick (3.0)
| | | | | | | | | referral_source = SVI
| | | | | | | | | FTI <= 151
| | | | | | | | | FTI <= 69
| | | | | | | | | | FTI <= 61: negative (2.08/0.08)
| | | | | | | | | | FTI > 61: sick (6.22)
| | | | | | | | | | FTI > 69: sick (148.36)
| | | | | | | | | | FTI > 151
| | | | | | | | | | | age <= 73: negative (5.28/1.28)
| | | | | | | | | | | age > 73: sick (4.06)
| | | | | | | | | | | referral_source = STMW: sick (0.0)
| | | | | | | | | | | referral_source = SVHD: sick (3.0)
| | | | | | TT4 > 136
| | | | | | | T4U_measured = t: negative (12.0/1.0)
| | | | | | | T4U_measured = f: sick (2.0)
| | | | | | T3_measured = f: negative (79.64/0.31)

T3 > 1.1
| FTI <= 171
| | TT4 <= 56
| | | referral_source = SVHC: negative (3.89)
| | | referral_source = other: negative (32.25/0.9)
| | | referral_source = SVI
| | | | TSH <= 14: sick (7.79/1.0)
| | | | | TSH > 14: negative (4.98)
| | | | | referral_source = STMW: negative (2.03)
| | | | | referral_source = SVHD: negative (0.01)
| | | | TT4 > 56: negative (3187.63/1.0)
| | FTI > 171
| | | T3 <= 2.6
| | | | on_thyroxine = f

```

```

|   |   |   |   T3_measured = t
|   |   |   |   TSH <= 0.27: sick (19.02/3.02)
|   |   |   |   TSH > 0.27: negative (11.76)
|   |   |   T3_measured = f: negative (8.79)
|   |   on_thyroxine = t: negative (19.64)
|   |   T3 > 2.6: negative (83.58)

Number of Leaves : 34

Size of the tree : 61

Time taken to build model: 0.14 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3727          98.807 %
Incorrectly Classified Instances    45           1.193 %
Kappa statistic                   0.8943
Mean absolute error               0.0146
Root mean squared error           0.1054
Relative absolute error           12.685 %
Root relative squared error      43.9447 %
Total Number of Instances         3772

==== Detailed Accuracy By Class ====

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC     ROC Area  PRC Area  Class
      0.995   0.117    0.992     0.995   0.994     0.895   0.951    0.993    negative
      0.883   0.005    0.919     0.883   0.901     0.895   0.951    0.864    sick
Weighted Avg.      0.988   0.110    0.988     0.988   0.988     0.895   0.951    0.985

==== Confusion Matrix ====

      a     b  <-- classified as
3523  18 |     a = negative
      27  204 |     b = sick

```

### 2.5.3 Cross Validation (15 Folds)

```

==== Run information ====

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   15-fold cross-validation

==== Classifier model (full training set) ====

J48 pruned tree
-----
T3 <= 1.1

```

```

| T3_measured = t
|   TT4 <= 53
|     | TSH <= 19
|       |   FTI <= 78
|         |     |   T4U <= 0.67
|           |       |   T3 <= 0.5: sick (3.0/1.0)
|             |       |   T3 > 0.5: negative (2.0)
|               |       |   T4U > 0.67: sick (7.0)
|                 |       FTI > 78: negative (8.0)
|                   |       TSH > 19: negative (53.0)
|   TT4 > 53
|     | TT4 <= 136
|       |   referral_source = SVHC: sick (9.0)
|         |   referral_source = other
|           |     |   sick = f
|             |       |   age <= 77
|               |       |       |   query_hypothyroid = f
|                 |       |       |       |   on_thyroxine = f
|                   |       |       |       |       |   T3 <= 0.5: sick (5.0/1.0)
|                     |       |       |       |       |   T3 > 0.5: negative (21.0/1.0)
|                       |       |       |       |       |   on_thyroxine = t: sick (5.0/2.0)
|                         |       |       |       |       |   query_hypothyroid = t: sick (4.0/1.0)
|                           |       |       |       |       |       |   age > 77: sick (8.0)
|                             |       |       |       |       |       |   sick = t: sick (3.0)
|                               |       |       |       |       |       |   referral_source = SVI
|                                 |       |       |       |       |       |       |   FTI <= 151
|                                   |       |       |       |       |       |       |       |   FTI <= 69
|                                     |       |       |       |       |       |       |       |       |   FTI <= 61: negative (2.08/0.08)
|                                       |       |       |       |       |       |       |       |       |   FTI > 61: sick (6.22)
|                                         |       |       |       |       |       |       |       |       |   FTI > 69: sick (148.36)
|                                           |       |       |       |       |       |       |       |       |   FTI > 151
|                                             |       |       |       |       |       |       |       |       |       |   age <= 73: negative (5.28/1.28)
|                                               |       |       |       |       |       |       |       |       |       |       |   age > 73: sick (4.06)
|                                                 |       |       |       |       |       |       |       |       |       |       |   referral_source = STMW: sick (0.0)
|                                                   |       |       |       |       |       |       |       |       |       |       |   referral_source = SVHD: sick (3.0)
|   TT4 > 136
|     |   T4U_measured = t: negative (12.0/1.0)
|       |   T4U_measured = f: sick (2.0)
|   T3_measured = f: negative (79.64/0.31)

T3 > 1.1
|   FTI <= 171
|     |   TT4 <= 56
|       |     |   referral_source = SVHC: negative (3.89)
|         |     |   referral_source = other: negative (32.25/0.9)
|           |     |   referral_source = SVI
|             |     |       |   TSH <= 14: sick (7.79/1.0)
|               |     |       |       |   TSH > 14: negative (4.98)
|                 |     |       |       |   referral_source = STMW: negative (2.03)
|                   |     |       |       |   referral_source = SVHD: negative (0.01)
|                     |     |       |       |   TT4 > 56: negative (3187.63/1.0)
|           |     |   FTI > 171
|             |       |   T3 <= 2.6
|               |       |       |   on_thyroxine = f
|                 |       |       |       |   T3_measured = t
|                   |       |       |       |       |   TSH <= 0.27: sick (19.02/3.02)
|                     |       |       |       |       |       |   TSH > 0.27: negative (11.76)
|                       |       |       |       |       |       |   T3_measured = f: negative (8.79)
|                         |       |       |       |       |       |       |   on_thyroxine = t: negative (19.64)
|                           |       |       |       |       |       |       |       |   T3 > 2.6: negative (83.58)

Number of Leaves : 34
Size of the tree : 61

Time taken to build model: 0.11 seconds

==== Stratified cross-validation ====
==== Summary ====

  Correctly Classified Instances      3721          98.6479 %
  Incorrectly Classified Instances     51           1.3521 %
  Kappa statistic                      0.8802
  Mean absolute error                  0.0166
  Root mean squared error              0.1116
  Relative absolute error              14.3774 %
  Root relative squared error          46.5532 %
  Total Number of Instances            3772

==== Detailed Accuracy By Class ====

          TP Rate  FP Rate  Precision  Recall   F-Measure  MCC    ROC Area  PRC Area  Class
          0.994    0.130    0.992    0.994    0.993    0.880    0.964    0.994    negative
          0.870    0.006    0.905    0.870    0.887    0.880    0.964    0.832    sick
  Weighted Avg.    0.986    0.122    0.986    0.986    0.986    0.880    0.964    0.984

==== Confusion Matrix ====

```

a	b	<-- classified as
3520	21	a = negative
30	201	b = sick

## 2.5.4 Percentage Split 50%

```
==== Run information ====
Scheme: weka.classifiers.trees.J48 -C 0.25 -M
Relation: sick
Instances: 3772
Attributes: 30
age
sex
on_thyroxine
query_on_thyroxine
on_antithyroid_medication
sick
pregnant
thyroid_surgery
I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH
T3_measured
T3
TT4_measured
TT4
T4U_measured
T4U
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class

Test mode: split 50.0% train, remainder test
```

```

|   |   |   |   referral_source = SVHD: sick (3.0)
|   |   |   TT4 > 136
|   |   |   |   T4U_measured = t: negative (12.0/1.0)
|   |   |   |   T4U_measured = f: sick (2.0)
|   |   T3_measured = f: negative (79.64/0.31)
T3 > 1.1
|   FTI <= 171
|   |   TT4 <= 56
|   |   |   referral_source = SVHC: negative (3.89)
|   |   |   referral_source = other: negative (32.25/0.9)
|   |   |   referral_source = SVI
|   |   |   |   TSH <= 14: sick (7.79/1.0)
|   |   |   |   TSH > 14: negative (4.98)
|   |   |   referral_source = STMW: negative (2.03)
|   |   |   referral_source = SVHD: negative (0.01)
|   |   TT4 > 56: negative (3187.63/1.0)
FTI > 171
|   |   T3 <= 2.6
|   |   |   on_thyroxine = f
|   |   |   |   T3_measured = t
|   |   |   |   |   TSH <= 0.27: sick (19.02/3.02)
|   |   |   |   |   TSH > 0.27: negative (11.76)
|   |   |   |   T3_measured = f: negative (8.79)
|   |   |   on_thyroxine = t: negative (19.64)
|   |   T3 > 2.6: negative (83.58)

Number of Leaves : 34
Size of the tree : 61

Time taken to build model: 0.12 seconds
==== Evaluation on test split ====
Time taken to test model on test split: 0.01 seconds
==== Summary ====
Correctly Classified Instances      1855          98.3563 %
Incorrectly Classified Instances    31           1.6437 %
Kappa statistic                      0.8523
Mean absolute error                  0.0236
Root mean squared error              0.1227
Relative absolute error              20.428 %
Root relative squared error         50.8528 %
Total Number of Instances           1886

==== Detailed Accuracy By Class ====


|               | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC   | ROC Area | PRC Area | Class |
|---------------|---------|---------|-----------|--------|-----------|-------|----------|----------|-------|
| 0.994         | 0.179   | 0.988   | 0.994     | 0.991  | 0.853     | 0.905 | 0.986    | negative |       |
| 0.821         | 0.006   | 0.906   | 0.821     | 0.861  | 0.853     | 0.905 | 0.795    | sick     |       |
| Weighted Avg. | 0.984   | 0.169   | 0.983     | 0.984  | 0.983     | 0.853 | 0.905    | 0.974    |       |


==== Confusion Matrix ====


| a    | b  | <-- classified as |
|------|----|-------------------|
| 1759 | 10 | a = negative      |
| 21   | 96 | b = sick          |


```

## 2.5.5 Percentage Split 60%

```

==== Run information ====
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:     sick
Instances:    3772
Attributes:   30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych

```

```

TSH_measured
TSH
T3_measured
T3
TT4_measured
TT4
T4U_measured
T4U
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Test mode: split 60.0% train, remainder test

==== Classifier model (full training set) ====

J48 pruned tree
-----
T3 <= 1.1
| T3_measured = t
| | TT4 <= 53
| | | TSH <= 19
| | | | FTI <= 78
| | | | | T4U <= 0.67
| | | | | | T3 <= 0.5: sick (3.0/1.0)
| | | | | | T3 > 0.5: negative (2.0)
| | | | | | T4U > 0.67: sick (7.0)
| | | | | | FTI > 78: negative (8.0)
| | | | | | TSH > 19: negative (53.0)
| | | | TT4 > 53
| | | | | TT4 <= 136
| | | | | | referral_source = SVHC: sick (9.0)
| | | | | | referral_source = other
| | | | | | sick = f
| | | | | | age <= 77
| | | | | | | query_hypothyroid = f
| | | | | | | | on_thyroxine = f
| | | | | | | | | T3 <= 0.5: sick (5.0/1.0)
| | | | | | | | | T3 > 0.5: negative (21.0/1.0)
| | | | | | | | | on_thyroxine = t: sick (5.0/2.0)
| | | | | | | | | query_hypothyroid = t: sick (4.0/1.0)
| | | | | | | | | age > 77: sick (8.0)
| | | | | | | | | sick = t: sick (3.0)
| | | | | | | | | referral_source = SVI
| | | | | | | | | FTI <= 151
| | | | | | | | | FTI <= 69
| | | | | | | | | | FTI <= 61: negative (2.08/0.08)
| | | | | | | | | | FTI > 61: sick (6.22)
| | | | | | | | | | FTI > 69: sick (148.36)
| | | | | | | | | | FTI > 151
| | | | | | | | | | age <= 73: negative (5.28/1.28)
| | | | | | | | | | age > 73: sick (4.06)
| | | | | | | | | | referral_source = STMW: sick (0.0)
| | | | | | | | | | referral_source = SVHD: sick (3.0)
| | | | | | TT4 > 136
| | | | | | | T4U_measured = t: negative (12.0/1.0)
| | | | | | | T4U_measured = f: sick (2.0)
| | | | | | T3_measured = f: negative (79.64/0.31)

T3 > 1.1
| FTI <= 171
| | TT4 <= 56
| | | referral_source = SVHC: negative (3.89)
| | | | referral_source = other: negative (32.25/0.9)
| | | | referral_source = SVI
| | | | | TSH <= 14: sick (7.79/1.0)
| | | | | TSH > 14: negative (4.98)
| | | | | referral_source = STMW: negative (2.03)
| | | | | referral_source = SVHD: negative (0.01)
| | | | TT4 > 56: negative (3187.63/1.0)
| | FTI > 171
| | | T3 <= 2.6
| | | | on_thyroxine = f
| | | | | T3_measured = t
| | | | | | TSH <= 0.27: sick (19.02/3.02)
| | | | | | TSH > 0.27: negative (11.76)
| | | | | | T3_measured = f: negative (8.79)
| | | | | | on_thyroxine = t: negative (19.64)
| | | | | | T3 > 2.6: negative (83.58)

Number of Leaves : 34
Size of the tree : 61

```

```

Time taken to build model: 0.13 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

==== Summary ===

Correctly Classified Instances      1488          98.6083 %
Incorrectly Classified Instances   21           1.3917 %
Kappa statistic                   0.8623
Mean absolute error               0.0215
Root mean squared error          0.1139
Relative absolute error          18.7672 %
Root relative squared error     48.3365 %
Total Number of Instances        1509

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC    ROC Area   PRC Area   Class
0.999     0.213     0.987     0.999     0.993     0.868     0.944     0.993   negative
0.787     0.001     0.972     0.787     0.870     0.868     0.944     0.857   sick
Weighted Avg.   0.986     0.201     0.986     0.986     0.985     0.868     0.944     0.985

==== Confusion Matrix ====

      a      b  <-- classified as
1418   2 |    a = negative
      19  70 |    b = sick

```

## 2.5.6 Percentage Split 66%

```

==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 66.0% train, remainder test

==== Classifier model (full training set) ===

J48 pruned tree
-----
T3 <= 1.1
|  T3_measured = t
|  |  TT4 <= 53
|  |  |  TSH <= 19
|  |  |  |  FTI <= 78
|  |  |  |  |  T4U <= 0.67
|  |  |  |  |  |  T3 <= 0.5: sick (3.0/1.0)
|  |  |  |  |  |  |  T3 > 0.5: negative (2.0)
|  |  |  |  |  |  |  T4U > 0.67: sick (7.0)
|  |  |  |  |  |  |  FTI > 78: negative (8.0)

```

```

|   |   |   TSH > 19: negative (53.0)
|   |   TT4 > 53
|   |   |   TT4 <= 136
|   |   |   |   referral_source = SVHC: sick (9.0)
|   |   |   |   referral_source = other
|   |   |   |   |   sick = f
|   |   |   |   |   age <= 77
|   |   |   |   |   |   query_hypothyroid = f
|   |   |   |   |   |   |   on_thyroxine = f
|   |   |   |   |   |   |   |   T3 <= 0.5: sick (5.0/1.0)
|   |   |   |   |   |   |   |   T3 > 0.5: negative (21.0/1.0)
|   |   |   |   |   |   |   on_thyroxine = t: sick (5.0/2.0)
|   |   |   |   |   |   |   query_hypothyroid = t: sick (4.0/1.0)
|   |   |   |   |   |   age > 77: sick (8.0)
|   |   |   |   |   |   sick = t: sick (3.0)
|   |   |   |   |   |   referral_source = SVI
|   |   |   |   |   |   FTI <= 151
|   |   |   |   |   |   |   FTI <= 69
|   |   |   |   |   |   |   |   FTI <= 61: negative (2.08/0.08)
|   |   |   |   |   |   |   |   FTI > 61: sick (6.22)
|   |   |   |   |   |   |   |   FTI > 69: sick (148.36)
|   |   |   |   |   |   |   FTI > 151
|   |   |   |   |   |   |   |   age <= 73: negative (5.28/1.28)
|   |   |   |   |   |   |   |   age > 73: sick (4.06)
|   |   |   |   |   |   |   |   referral_source = STMW: sick (0.0)
|   |   |   |   |   |   |   |   referral_source = SVHD: sick (3.0)
|   |   |   |   |   |   TT4 > 136
|   |   |   |   |   |   T4U_measured = t: negative (12.0/1.0)
|   |   |   |   |   |   T4U_measured = f: sick (2.0)
|   |   |   |   |   T3_measured = f: negative (79.64/0.31)

T3 > 1.1
|   FTI <= 171
|   |   TT4 <= 56
|   |   |   referral_source = SVHC: negative (3.89)
|   |   |   referral_source = other: negative (32.25/0.9)
|   |   |   referral_source = SVI
|   |   |   |   TSH <= 14: sick (7.79/1.0)
|   |   |   |   TSH > 14: negative (4.98)
|   |   |   |   referral_source = STMW: negative (2.03)
|   |   |   |   referral_source = SVHD: negative (0.01)
|   |   |   TT4 > 56: negative (3187.63/1.0)
|   FTI > 171
|   |   T3 <= 2.6
|   |   |   on_thyroxine = f
|   |   |   |   T3_measured = t
|   |   |   |   |   TSH <= 0.27: sick (19.02/3.02)
|   |   |   |   |   TSH > 0.27: negative (11.76)
|   |   |   |   |   T3_measured = f: negative (8.79)
|   |   |   |   on_thyroxine = t: negative (19.64)
|   |   |   T3 > 2.6: negative (83.58)

Number of Leaves : 34
Size of the tree : 61

Time taken to build model: 0.11 seconds
== Evaluation on test split ==
Time taken to test model on test split: 0.03 seconds
== Summary ==
Correctly Classified Instances      1257      98.0499 %
Incorrectly Classified Instances     25       1.9501 %
Kappa statistic                      0.8149
Mean absolute error                  0.0234
Root mean squared error              0.1336
Relative absolute error              20.4604 %
Root relative squared error          56.914  %
Total Number of Instances            1282

== Detailed Accuracy By Class ==
      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
          0.993    0.213    0.987    0.993    0.990    0.816    0.878    0.979  negative
          0.787    0.007    0.868    0.787    0.825    0.816    0.878    0.809    sick
Weighted Avg.      0.980    0.201    0.980    0.980    0.980    0.816    0.878    0.969

== Confusion Matrix ==
      a     b  <- classified as
1198   9 |     a = negative
     16   59 |     b = sick

```

### 2.5.7 Leave one out Fold

```

==== Run information ====

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   3772-fold cross-validation

==== Classifier model (full training set) ====
J48 pruned tree
-----
T3 <= 1.1
|  T3_measured = t
|  |  TT4 <= 53
|  |  |  TSH <= 19
|  |  |  |  FTI <= 78
|  |  |  |  |  T4U <= 0.67
|  |  |  |  |  |  T3 <= 0.5: sick (3.0/1.0)
|  |  |  |  |  |  T3 > 0.5: negative (2.0)
|  |  |  |  |  |  T4U > 0.67: sick (7.0)
|  |  |  |  |  FTI > 78: negative (8.0)
|  |  |  |  TSH > 19: negative (53.0)
|  |  TT4 > 53
|  |  |  TT4 <= 136
|  |  |  |  referral_source = SVHC: sick (9.0)
|  |  |  |  referral_source = other
|  |  |  |  |  sick = f
|  |  |  |  |  |  age <= 77
|  |  |  |  |  |  |  query_hypothyroid = f
|  |  |  |  |  |  |  |  on_thyroxine = f
|  |  |  |  |  |  |  |  |  T3 <= 0.5: sick (5.0/1.0)
|  |  |  |  |  |  |  |  |  T3 > 0.5: negative (21.0/1.0)
|  |  |  |  |  |  |  |  |  on_thyroxine = t: sick (5.0/2.0)
|  |  |  |  |  |  |  |  |  query_hypothyroid = t: sick (4.0/1.0)
|  |  |  |  |  |  |  |  |  age > 77: sick (8.0)
|  |  |  |  |  |  |  |  |  sick = t: sick (3.0)
|  |  |  |  |  |  |  |  |  referral_source = SVI
|  |  |  |  |  |  |  |  |  |  FTI <= 151
|  |  |  |  |  |  |  |  |  |  FTI <= 69: negative (2.08/0.08)
|  |  |  |  |  |  |  |  |  |  FTI > 69: sick (148.36)
|  |  |  |  |  |  |  |  |  |  FTI > 151
|  |  |  |  |  |  |  |  |  |  age <= 73: negative (5.28/1.28)
|  |  |  |  |  |  |  |  |  |  age > 73: sick (4.06)
|  |  |  |  |  |  |  |  |  |  referral_source = STMW: sick (0.0)
|  |  |  |  |  |  |  |  |  |  referral_source = SVHD: sick (3.0)
|  |  |  |  |  TT4 > 136
|  |  |  |  |  |  T4U_measured = t: negative (12.0/1.0)
|  |  |  |  |  |  |  T4U_measured = f: sick (2.0)
|  |  |  |  |  T3_measured = f: negative (79.64/0.31)
T3 > 1.1
|  FTI <= 171

```

```

|   |   TT4 <= 56
|   |   | referral_source = SVHC: negative (3.89)
|   |   | referral_source = other: negative (32.25/0.9)
|   |   | referral_source = SVI
|   |   |   | TSH <= 14: sick (7.79/1.0)
|   |   |   | TSH > 14: negative (4.98)
|   |   |   | referral_source = STMW: negative (2.03)
|   |   |   | referral_source = SVHD: negative (0.01)
|   |   TT4 > 56: negative (3187.63/1.0)
|   FTI > 171
|   |   T3 <= 2.6
|   |   | on_thyroxine = f
|   |   |   | T3_measured = t
|   |   |   |   | TSH <= 0.27: sick (19.02/3.02)
|   |   |   |   | TSH > 0.27: negative (11.76)
|   |   |   |   | T3_measured = f: negative (8.79)
|   |   |   | on_thyroxine = t: negative (19.64)
|   |   T3 > 2.6: negative (83.58)

Number of Leaves : 34
Size of the tree : 61

Time taken to build model: 0.12 seconds

== Stratified cross-validation ==
== Summary ==

Correctly Classified Instances      3716          98.5154 %
Incorrectly Classified Instances     56           1.4846 %
Kappa statistic                      0.8682
Mean absolute error                  0.0162
Root mean squared error              0.1154
Relative absolute error              14.044 %
Root relative squared error         48.1062 %
Total Number of Instances           3772

== Detailed Accuracy By Class ==

      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC    ROC Area  PRC Area  Class
      0.994    0.143    0.991    0.994    0.992    0.868    0.963    0.994    negative
      0.857    0.006    0.896    0.857    0.876    0.868    0.963    0.905    sick
Weighted Avg.      0.985    0.135    0.985    0.985    0.985    0.868    0.963    0.989

== Confusion Matrix ==

      a     b  <- classified as
3518  23 |  a = negative
      33 198 |  b = sick

```

### 2.5.8 Conclusion

C4.5			
Evaluation Process		Correctly Classified Instances	notes
Cross Validation	5 folds	98.4358%	
	10 folds	98.8070%	BEST
	15 folds	98.6479%	
Percentage Split	50.00%	98.3563%	
	60.00%	98.6083%	
	66.00%	98.0499%	
Leave One Out Fold		98.5154%	

## 2.6 PART "default settings"

### 2.6.1 Cross Validation (5 Folds)

```

== Run information ==
Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:     sick

```

```
Instances:      3772
Attributes:    30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:     5-fold cross-validation

==== Classifier model (full training set) ====

PART decision list
-----

T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3187.63/1.0)

referral_source = other AND
T3_measured = f: negative (112.3/1.0)

referral_source = other AND
sick = f AND
T3 <= 2.6 AND
TT4 <= 59: negative (54.58)

T3 > 1.8 AND
T3 > 2.6: negative (68.81/0.33)

TT4 <= 50 AND
TSH > 20: negative (28.11)

referral_source = SVI AND
T3_measured = t AND
FTI <= 157 AND
TT4 > 53 AND
FTI > 65: sick (156.14/0.5)

sick = f AND
T4U_measured = f: negative (12.28/1.2)

sick = f AND
TT4 > 193: negative (11.67)

sick = f AND
referral_source = SVHC: sick (10.04/2.04)

sick = f AND
on_thyroxine = t AND
TSH <= 1.2: negative (12.09)

sick = f AND
on_thyroxine = f AND
referral_source = other AND
age <= 77 AND
T3 <= 1.7 AND
T3 > 0.5 AND
query_hypothyroid = f: negative (22.0/2.0)

referral_source = other AND
query_hypothyroid = f AND
T4U > 0.8: sick (21.0)
```

```

sick = f AND
referral_source = other AND
TSH > 0.22: sick (8.0/1.0)

sick = f AND
referral_source = SVI AND
T3 <= 1.9 AND
FTI > 78: negative (23.05/4.0)

referral_source = SVI AND
T3 > 1: sick (12.95/0.3)

referral_source = SVI AND
sex = M: sick (10.48/2.28)

sick = f AND
referral_source = SVI AND
age > 51: negative (7.16/1.0)

referral_source = other: negative (4.0/1.0)

T3 <= 1.5: sick (7.56/0.56)

: negative (2.16)

Number of Rules : 20

Time taken to build model: 0.17 seconds

==== Stratified cross-validation ====
==== Summary ===

    Correctly Classified Instances      3708           98.3033 %
    Incorrectly Classified Instances   64            1.6967 %
    Kappa statistic                   0.853
    Mean absolute error              0.0185
    Root mean squared error          0.1193
    Relative absolute error          16.072 %
    Root relative squared error     49.7476 %
    Total Number of Instances        3772

==== Detailed Accuracy By Class ===

          TP Rate   FP Rate   Precision   Recall   F-Measure   MCC     ROC Area   PRC Area   Class
          0.991     0.134     0.991     0.991     0.991     0.853     0.972     0.996     negative
          0.866     0.009     0.858     0.866     0.862     0.853     0.972     0.873     sick
Weighted Avg.   0.983     0.127     0.983     0.983     0.983     0.853     0.972     0.989

==== Confusion Matrix ===

      a      b  <- classified as
3508  33 |  a = negative
      31  200 |  b = sick

```

## 2.6.2 Cross Validation (10 Folds)

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick
Instances:   3772
Attributes:  30
             age
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             I131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary
             psych
             TSH_measured
             TSH
             T3_measured
             T3
             TT4_measured

```

```

TT4
T4U_measured
T4U
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Test mode: 10-fold cross-validation

==== Classifier model (full training set) ===

PART decision list
-----

T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3187.63/1.0)

referral_source = other AND
T3_measured = f: negative (112.3/1.0)

referral_source = other AND
sick = f AND
T3 <= 2.6 AND
TT4 <= 59: negative (54.58)

T3 > 1.8 AND
T3 > 2.6: negative (68.81/0.33)

TT4 <= 50 AND
TSH > 20: negative (28.11)

referral_source = SVI AND
T3_measured = t AND
FTI <= 157 AND
TT4 > 53 AND
FTI > 65: sick (156.14/0.5)

sick = f AND
T4U_measured = f: negative (12.28/1.2)

sick = f AND
TT4 > 193: negative (11.67)

sick = f AND
referral_source = SVHC: sick (10.04/2.04)

sick = f AND
on_thyroxine = t AND
TSH <= 1.2: negative (12.09)

sick = f AND
on_thyroxine = f AND
referral_source = other AND
age <= 77 AND
T3 <= 1.7 AND
T3 > 0.5 AND
query_hypothyroid = f: negative (22.0/2.0)

referral_source = other AND
query_hypothyroid = f AND
T4U > 0.8: sick (21.0)

sick = f AND
referral_source = other AND
TSH > 0.22: sick (8.0/1.0)

sick = f AND
referral_source = SVI AND
T3 <= 1.9 AND
FTI > 78: negative (23.05/4.0)

referral_source = SVI AND
T3 > 1: sick (12.95/0.3)

referral_source = SVI AND
sex = M: sick (10.48/2.28)

sick = f AND
referral_source = SVI AND
age > 51: negative (7.16/1.0)

referral_source = other: negative (4.0/1.0)

T3 <= 1.5: sick (7.56/0.56)

```

```

: negative (2.16)

Number of Rules : 20

Time taken to build model: 0.15 seconds

==== Stratified cross-validation ====
==== Summary ====

    Correctly Classified Instances      3720           98.6214 %
    Incorrectly Classified Instances   52            1.3786 %
    Kappa statistic                   0.8801
    Mean absolute error              0.0164
    Root mean squared error          0.1107
    Relative absolute error          14.2412 %
    Root relative squared error     46.1561 %
    Total Number of Instances       3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC     ROC Area   PRC Area   Class
      0.993     0.113     0.993     0.993     0.993     0.880     0.956     0.993     negative
      0.887     0.007     0.887     0.887     0.887     0.880     0.956     0.848     sick
Weighted Avg.   0.986     0.106     0.986     0.986     0.986     0.880     0.956     0.984

==== Confusion Matrix ====

      a     b  <-- classified as
  3515  26 |  a = negative
    26  205 |  b = sick

```

### 2.6.3 Cross Validation (15 Folds)

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   15-fold cross-validation

==== Classifier model (full training set) ====

PART decision list
-----
T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3187.63/1.0)

referral_source = other AND
T3_measured = f: negative (112.3/1.0)

```

```

referral_source = other AND
sick = f AND
T3 <= 2.6 AND
TT4 <= 59: negative (54.58)

T3 > 1.8 AND
T3 > 2.6: negative (68.81/0.33)

TT4 <= 50 AND
TSH > 20: negative (28.11)

referral_source = SVI AND
T3_measured = t AND
FTI <= 157 AND
TT4 > 53 AND
FTI > 65: sick (156.14/0.5)

sick = f AND
T4U_measured = f: negative (12.28/1.2)

sick = f AND
TT4 > 193: negative (11.67)

sick = f AND
referral_source = SVHC: sick (10.04/2.04)

sick = f AND
on_thyroxine = t AND
TSH <= 1.2: negative (12.09)

sick = f AND
on_thyroxine = f AND
referral_source = other AND
age <= 77 AND
T3 <= 1.7 AND
T3 > 0.5 AND
query_hypothyroid = f: negative (22.0/2.0)

referral_source = other AND
query_hypothyroid = f AND
T4U > 0.8: sick (21.0)

sick = f AND
referral_source = other AND
TSH > 0.22: sick (8.0/1.0)

sick = f AND
referral_source = SVI AND
T3 <= 1.9 AND
FTI > 78: negative (23.05/4.0)

referral_source = SVI AND
T3 > 1: sick (12.95/0.3)

referral_source = SVI AND
sex = M: sick (10.48/2.28)

sick = f AND
referral_source = SVI AND
age > 51: negative (7.16/1.0)

referral_source = other: negative (4.0/1.0)

T3 <= 1.5: sick (7.56/0.56)

: negative (2.16)

Number of Rules : 20

Time taken to build model: 0.15 seconds

==== Stratified cross-validation ===
==== Summary ===

Correctly Classified Instances      3719          98.5949 %
Incorrectly Classified Instances    53           1.4051 %
Kappa statistic                   0.8785
Mean absolute error               0.0169
Root mean squared error           0.1134
Relative absolute error            14.7071 %
Root relative squared error       47.2988 %
Total Number of Instances         3772

==== Detailed Accuracy By Class ===

TP Rate   FP Rate   Precision   Recall   F-Measure   MCC
ROC Area   PRC Area   Class

```

```

0.992    0.108    0.993    0.992    0.993    0.879    0.947    0.989    negative
0.892    0.008    0.880    0.892    0.886    0.879    0.947    0.839    sick
Weighted Avg.   0.986    0.102    0.986    0.986    0.986    0.879    0.947    0.980

==== Confusion Matrix ====
a      b    <-- classified as
3513  28 |     a = negative
 25  206 |     b = sick

```

## 2.6.4 Percentage Split 50%

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 50.0% train, remainder test

==== Classifier model (full training set) ===

PART decision list
-----
T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3187.63/1.0)

referral_source = other AND
T3_measured = f: negative (112.3/1.0)

referral_source = other AND
sick = f AND
T3 <= 2.6 AND
TT4 <= 59: negative (54.58)

T3 > 1.8 AND
T3 > 2.6: negative (68.81/0.33)

TT4 <= 50 AND
TSH > 20: negative (28.11)

referral_source = SVI AND
T3_measured = t AND
FTI <= 157 AND
TT4 > 53 AND
FTI > 65: sick (156.14/0.5)

sick = f AND
T4U_measured = f: negative (12.28/1.2)

sick = f AND
TT4 > 193: negative (11.67)

```

```

sick = f AND
referral_source = SVHC: sick (10.04/2.04)

sick = f AND
on_thyroxine = t AND
TSH <= 1.2: negative (12.09)

sick = f AND
on_thyroxine = f AND
referral_source = other AND
age <= 77 AND
T3 <= 1.7 AND
T3 > 0.5 AND
query_hypothyroid = f: negative (22.0/2.0)

referral_source = other AND
query_hypothyroid = f AND
T4U > 0.8: sick (21.0)

sick = f AND
referral_source = other AND
TSH > 0.22: sick (8.0/1.0)

sick = f AND
referral_source = SVI AND
T3 <= 1.9 AND
FTI > 78: negative (23.05/4.0)

referral_source = SVI AND
T3 > 1: sick (12.95/0.3)

referral_source = SVI AND
sex = M: sick (10.48/2.28)

sick = f AND
referral_source = SVI AND
age > 51: negative (7.16/1.0)

referral_source = other: negative (4.0/1.0)

T3 <= 1.5: sick (7.56/0.56)

: negative (2.16)

Number of Rules : 20

Time taken to build model: 0.14 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

==== Summary ===

Correctly Classified Instances      1860          98.6214 %
Incorrectly Classified Instances   26           1.3786 %
Kappa statistic                   0.8806
Mean absolute error               0.0162
Root mean squared error          0.1114
Relative absolute error          14.0052 %
Root relative squared error     46.1727 %
Total Number of Instances        1886

==== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
      0.993    0.120    0.992    0.993    0.993    0.881    0.964    0.995    negative
      0.880    0.007    0.896    0.880    0.888    0.881    0.964    0.899    sick
Weighted Avg.       0.986    0.113    0.986    0.986    0.986    0.881    0.964    0.989

==== Confusion Matrix ===

      a      b  <- classified as
1757   12 |      a = negative
      14   103 |      b = sick

```

## 2.6.5 Percentage Split 60%

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick
Instances:   3772
Attributes:  30

```

```
age
sex
on_thyroxine
query_on_thyroxine
on_antithyroid_medication
sick
pregnant
thyroid_surgery
I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH
T3_measured
T3
TT4_measured
TT4
T4U_measured
T4U
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Test mode: split 60.0% train, remainder test

==== Classifier model (full training set) ====

PART decision list
-----
T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3187.63/1.0)

referral_source = other AND
T3_measured = f: negative (112.3/1.0)

referral_source = other AND
sick = f AND
T3 <= 2.6 AND
TT4 <= 59: negative (54.58)

T3 > 1.8 AND
T3 > 2.6: negative (68.81/0.33)

TT4 <= 50 AND
TSH > 20: negative (28.11)

referral_source = SVI AND
T3_measured = t AND
FTI <= 157 AND
TT4 > 53 AND
FTI > 65: sick (156.14/0.5)

sick = f AND
T4U_measured = f: negative (12.28/1.2)

sick = f AND
TT4 > 193: negative (11.67)

sick = f AND
referral_source = SVHC: sick (10.04/2.04)

sick = f AND
on_thyroxine = t AND
TSH <= 1.2: negative (12.09)

sick = f AND
on_thyroxine = f AND
referral_source = other AND
age <= 77 AND
T3 <= 1.7 AND
T3 > 0.5 AND
query_hypothyroid = f: negative (22.0/2.0)

referral_source = other AND
query_hypothyroid = f AND
T4U > 0.8: sick (21.0)

sick = f AND
```

```

referral_source = other AND
TSH > 0.22: sick (8.0/1.0)

sick = f AND
referral_source = SVI AND
T3 <= 1.9 AND
FTI > 78: negative (23.05/4.0)

referral_source = SVI AND
T3 > 1: sick (12.95/0.3)

referral_source = SVI AND
sex = M: sick (10.48/2.28)

sick = f AND
referral_source = SVI AND
age > 51: negative (7.16/1.0)

referral_source = other: negative (4.0/1.0)

T3 <= 1.5: sick (7.56/0.56)

: negative (2.16)

Number of Rules : 20

Time taken to build model: 0.16 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

==== Summary ===

Correctly Classified Instances      1476          97.8131 %
Incorrectly Classified Instances   33           2.1869 %
Kappa statistic                   0.7998
Mean absolute error               0.0245
Root mean squared error          0.1376
Relative absolute error          21.327 %
Root relative squared error     58.4131 %
Total Number of Instances        1509

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC    ROC Area   PRC Area   Class
      0.989     0.202     0.987     0.989     0.988     0.800     0.974     0.996   negative
      0.798     0.011     0.826     0.798     0.811     0.800     0.974     0.895   sick
Weighted Avg.      0.978     0.191     0.978     0.978     0.978     0.800     0.974     0.990

==== Confusion Matrix ===

      a     b  <- classified as
1405  15 |    a = negative
      18  71 |    b = sick

```

## 2.6.6 Percentage Split 66%

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3

```

```
TT4_measured
TT4
T4U_measured
T4U
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Test mode: split 66.0% train, remainder test

==== Classifier model (full training set) ====

PART decision list
-----
T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3187.63/1.0)

referral_source = other AND
T3_measured = f: negative (112.3/1.0)

referral_source = other AND
sick = f AND
T3 <= 2.6 AND
TT4 <= 59: negative (54.58)

T3 > 1.8 AND
T3 > 2.6: negative (68.81/0.33)

TT4 <= 50 AND
TSH > 20: negative (28.11)

referral_source = SVI AND
T3_measured = t AND
FTI <= 157 AND
TT4 > 53 AND
FTI > 65: sick (156.14/0.5)

sick = f AND
T4U_measured = f: negative (12.28/1.2)

sick = f AND
TT4 > 193: negative (11.67)

sick = f AND
referral_source = SVHC: sick (10.04/2.04)

sick = f AND
on_thyroxine = t AND
TSH <= 1.2: negative (12.09)

sick = f AND
on_thyroxine = f AND
referral_source = other AND
age <= 77 AND
T3 <= 1.7 AND
T3 > 0.5 AND
query_hypothyroid = f: negative (22.0/2.0)

referral_source = other AND
query_hypothyroid = f AND
T4U > 0.8: sick (21.0)

sick = f AND
referral_source = other AND
TSH > 0.22: sick (8.0/1.0)

sick = f AND
referral_source = SVI AND
T3 <= 1.9 AND
FTI > 78: negative (23.05/4.0)

referral_source = SVI AND
T3 > 1: sick (12.95/0.3)

referral_source = SVI AND
sex = M: sick (10.48/2.28)

sick = f AND
referral_source = SVI AND
age > 51: negative (7.16/1.0)

referral_source = other: negative (4.0/1.0)
```

```

T3 <= 1.5: sick (7.56/0.56)
: negative (2.16)

Number of Rules : 20

Time taken to build model: 0.18 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

==== Summary ===

Correctly Classified Instances      1266           98.752 %
Incorrectly Classified Instances   16            1.248 %
Kappa statistic                   0.8853
Mean absolute error               0.0152
Root mean squared error          0.1022
Relative absolute error          13.3222 %
Root relative squared error     43.5456 %
Total Number of Instances        1282

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.994     0.120     0.993     0.994     0.993     0.885     0.971     0.995     negative
      0.880     0.006     0.904     0.880     0.892     0.885     0.971     0.906     sick
Weighted Avg.       0.988     0.113     0.987     0.988     0.987     0.885     0.971     0.990

==== Confusion Matrix ===

    a     b  <- classified as
1200    7 |    a = negative
    9    66 |    b = sick

```

## 2.6.7 Leave one out Fold

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   3772-fold cross-validation

==== Classifier model (full training set) ===

PART decision list
-----
T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3187.63/1.0)

```

```

referral_source = other AND
T3_measured = f: negative (112.3/1.0)

referral_source = other AND
sick = f AND
T3 <= 2.6 AND
TT4 <= 59: negative (54.58)

T3 > 1.8 AND
T3 > 2.6: negative (68.81/0.33)

TT4 <= 50 AND
TSH > 20: negative (28.11)

referral_source = SVI AND
T3_measured = t AND
FTI <= 157 AND
TT4 > 53 AND
FTI > 65: sick (156.14/0.5)

sick = f AND
T4U_measured = f: negative (12.28/1.2)

sick = f AND
TT4 > 193: negative (11.67)

sick = f AND
referral_source = SVHC: sick (10.04/2.04)

sick = f AND
on_thyroxine = t AND
TSH <= 1.2: negative (12.09)

sick = f AND
on_thyroxine = f AND
referral_source = other AND
age <= 77 AND
T3 <= 1.7 AND
T3 > 0.5 AND
query_hypothyroid = f: negative (22.0/2.0)

referral_source = other AND
query_hypothyroid = f AND
T4U > 0.8: sick (21.0)

sick = f AND
referral_source = other AND
TSH > 0.22: sick (8.0/1.0)

sick = f AND
referral_source = SVI AND
T3 <= 1.9 AND
FTI > 78: negative (23.05/4.0)

referral_source = SVI AND
T3 > 1: sick (12.95/0.3)

referral_source = SVI AND
sex = M: sick (10.48/2.28)

sick = f AND
referral_source = SVI AND
age > 51: negative (7.16/1.0)

referral_source = other: negative (4.0/1.0)

T3 <= 1.5: sick (7.56/0.56)

: negative (2.16)

Number of Rules : 20

Time taken to build model: 0.14 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3730          98.8865 %
Incorrectly Classified Instances   42           1.1135 %
Kappa statistic                   0.902
Mean absolute error               0.0148
Root mean squared error          0.1007
Relative absolute error          12.8165 %
Root relative squared error     41.9956 %
Total Number of Instances         3772

```

```
==== Detailed Accuracy By Class ====
      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
      0.995    0.104    0.993    0.995    0.994    0.902    0.950    0.992    negative
      0.896    0.005    0.920    0.896    0.908    0.902    0.950    0.900    sick
Weighted Avg.  0.989    0.098    0.989    0.989    0.989    0.902    0.950    0.986

==== Confusion Matrix ====
      a     b  <- classified as
3523  18 |     a = negative
  24  207 |     b = sick
```

## 2.6.8 Conclusion

PART		
Evaluation Process	Correctly Classified Instances	notes
Cross Validation	5 folds	98.3033%
	10 folds	98.6214%
	15 folds	98.5949%
Percentage Split	50.00%	98.6214%
	60.00%	97.8131%
	66.00%	98.7520%
Leave One Out Fold	98.8865%	BEST

## 2.7 Random Forest "default settings"

### 2.7.1 Cross Validation (5 Folds)

```
==== Run information ====
Scheme:      weka.classifiers.trees.RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   5-fold cross-validation

==== Classifier model (full training set) ====
RandomForest
```

```

Bagging with 100 iterations and base learner

weka.classifiers.trees.RandomTree -K 0 -M 1.0 -V 0.001 -S 1 -do-not-check-capabilities

Time taken to build model: 1.08 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3704          98.1972 %
Incorrectly Classified Instances   68           1.8028 %
Kappa statistic                   0.8271
Mean absolute error               0.0339
Root mean squared error          0.112
Relative absolute error           29.4375 %
Root relative squared error     46.7242 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC    ROC Area  PRC Area  Class
0.997    0.247    0.984    0.997    0.990    0.833    0.997    1.000    negative
0.753    0.003    0.941    0.753    0.837    0.833    0.997    0.961    sick
Weighted Avg.                     0.982    0.232    0.981    0.982    0.981    0.833    0.997    0.997

==== Confusion Matrix ====

      a      b  <-- classified as
3530   11 |      a = negative
      57  174 |      b = sick

```

## 2.7.2 Cross Validation (10 Folds)

```

==== Run information ====

Scheme:      weka.classifiers.trees.RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              t131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   10-fold cross-validation

==== Classifier model (full training set) ====

RandomForest

Bagging with 100 iterations and base learner

weka.classifiers.trees.RandomTree -K 0 -M 1.0 -V 0.001 -S 1 -do-not-check-capabilities

Time taken to build model: 1.04 seconds

==== Stratified cross-validation ====
==== Summary ====

```

```

Correctly Classified Instances      3715          98.4889 %
Incorrectly Classified Instances   57           1.5111 %
Kappa statistic                   0.8567
Mean absolute error               0.0323
Root mean squared error           0.1083
Relative absolute error            28.054 %
Root relative squared error       45.1626 %
Total Number of Instances         3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC     ROC Area   PRC Area   Class
      0.998    0.212     0.986      0.998    0.992     0.861    0.996     1.000    negative
      0.788    0.002     0.958      0.788    0.865     0.861    0.996     0.966    sick
Weighted Avg.      0.985    0.199     0.985      0.985    0.984     0.861    0.996     0.998

==== Confusion Matrix ====

      a      b  <-- classified as
3533    8 |  a = negative
  49  182 |  b = sick

```

### 2.7.3 Cross Validation (15 Folds)

```

==== Run information ===

Scheme:      weka.classifiers.trees.RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   15-fold cross-validation

==== Classifier model (full training set) ===

RandomForest

Bagging with 100 iterations and base learner

weka.classifiers.trees.RandomTree -K 0 -M 1.0 -V 0.001 -S 1 -do-not-check-capabilities

Time taken to build model: 1.1 seconds

==== Stratified cross-validation ===
==== Summary ===

Correctly Classified Instances      3706          98.2503 %
Incorrectly Classified Instances   66           1.7497 %
Kappa statistic                   0.8351
Mean absolute error               0.033
Root mean squared error           0.1102
Relative absolute error            28.6136 %
Root relative squared error       45.9577 %
Total Number of Instances         3772

```

```
==== Detailed Accuracy By Class ====
      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
      0.996    0.225    0.985    0.996    0.991    0.839    0.997    1.000    negative
      0.775    0.004    0.927    0.775    0.844    0.839    0.997    0.962    sick
Weighted Avg.  0.983    0.212    0.982    0.983    0.982    0.839    0.997    0.998

==== Confusion Matrix ====
      a     b  <- classified as
3527  14 |  a = negative
      52  179 |  b = sick
```

## 2.7.4 Percentage Split 50%

```
==== Run information ====
Scheme:      weka.classifiers.trees.RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 50.0% train, remainder test

==== Classifier model (full training set) ====
RandomForest
Bagging with 100 iterations and base learner
weka.classifiers.trees.RandomTree -K 0 -M 1.0 -V 0.001 -S 1 -do-not-check-capabilities
Time taken to build model: 1.13 seconds

==== Evaluation on test split ====
Time taken to test model on test split: 0.19 seconds

==== Summary ====
  Correctly Classified Instances      1845           97.8261 %
  Incorrectly Classified Instances     41            2.1739 %
  Kappa statistic                   0.7868
  Mean absolute error               0.0387
  Root mean squared error          0.1237
  Relative absolute error          33.5603 %
  Root relative squared error     51.2639 %
  Total Number of Instances        1886

==== Detailed Accuracy By Class ====
      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
      0.997    0.308    0.980    0.997    0.989    0.797    0.996    1.000    negative
      0.692    0.003    0.942    0.692    0.798    0.797    0.996    0.953    sick
Weighted Avg.  0.978    0.289    0.978    0.978    0.977    0.797    0.996    0.997
```

```
==== Confusion Matrix ====

    a      b  <-- classified as
1764    5 |     a = negative
    36   81 |     b = sick
```

### 2.7.5 Percentage Split 60%

```
==== Run information ===

Scheme:      weka.classifiers.trees.RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              i131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 60.0% train, remainder test

==== Classifier model (full training set) ===

RandomForest

Bagging with 100 iterations and base learner

weka.classifiers.trees.RandomTree -K 0 -M 1.0 -V 0.001 -S 1 -do-not-check-capabilities

Time taken to build model: 1.03 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0.08 seconds

==== Summary ===

Correctly Classified Instances      1481      98.1445 %
Incorrectly Classified Instances    28       1.8555 %
Kappa statistic                   0.8087
Mean absolute error               0.0349
Root mean squared error           0.1173
Relative absolute error            30.4683 %
Root relative squared error      49.7712 %
Total Number of Instances         1509

==== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
0.999    0.292    0.982    0.999    0.990    0.820    0.997    1.000    negative
0.708    0.001    0.969    0.708    0.818    0.820    0.997    0.959    sick
Weighted Avg.  0.981    0.275    0.981    0.981    0.980    0.820    0.997    0.997

==== Confusion Matrix ====

    a      b  <-- classified as
1418    2 |     a = negative
    26   63 |     b = sick
```

### 2.7.6 Percentage Split 66%

```

==== Run information ===

Scheme:      weka.classifiers.trees.RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 66.0% train, remainder test

==== Classifier model (full training set) ===

RandomForest

Bagging with 100 iterations and base learner

weka.classifiers.trees.RandomTree -K 0 -M 1.0 -V 0.001 -S 1 -do-not-check-capabilities

Time taken to build model: 1.09 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0.07 seconds

==== Summary ===

  Correctly Classified Instances      1254           97.8159 %
  Incorrectly Classified Instances     28            2.1841 %
  Kappa statistic                   0.7701
  Mean absolute error               0.0368
  Root mean squared error          0.122
  Relative absolute error          32.2487 %
  Root relative squared error      51.9629 %
  Total Number of Instances        1282

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.998     0.333     0.980      0.998     0.989      0.783     0.997      1.000     negative
      0.667     0.002     0.943      0.667     0.781      0.783     0.997     0.949     sick
  Weighted Avg.     0.978     0.314     0.978      0.976     0.976      0.783     0.997     0.997

==== Confusion Matrix ===

      a      b  <-- classified as
1204    3 |    a = negative
      25   50 |    b = sick

```

### 2.7.7 Leave one out Fold

### 2.7.8 Conclusion

Random Forest			
Evaluation Process		Correctly Classified Instances	notes
Cross Validation	5 folds	98.1972%	
	10 folds	98.4889%	BEST
	15 folds	98.2503%	
Percentage Split	50.00%	97.8261%	
	60.00%	98.1445%	
	66.00%	97.8159%	
Leave One Out Fold		98.4624%	

After the tests we find that both **PART** and **C4.5** are the best performers and therefore we will choose them for the rest of the tests. as we can see tree based classifiers perform best with this dataset.

## 2.8 General Conclusion

Correctly Classified Instances by Algorithm						
Evaluation Process	One Rule	Naive Bayes	KNN	C4.5	PART	Random Forest
Cross Validation	5 folds	96.3680%	92.6034%	96.2089%	98.4358%	98.3033%
	10 folds	96.3945%	92.6034%	96.1824%	98.8070%	98.6214%
	15 folds	96.3945%	92.8155%	96.0498%	98.6479%	98.5949%
Percentage Split	50.00%	96.6066%	94.2736%	95.5992%	98.3563%	98.6214%
	60.00%	96.6203%	93.4394%	95.8250%	98.6083%	97.8131%
	66.00%	96.8799%	93.1357%	95.8658%	98.0499%	97.8159%
Leave One Out Fold		96.1294%	92.8685%	96.1824%	98.5154%	98.8865%
AVG Algorithms		96.48%	93.11%	95.99%	98.49%	98.51%
Rank Algorithms		4	6	5	2	1
						3
						98.17%

## Chapter 3

# Testing with preprocessed dataset and PART Algorithm

### 3.1 Non supervised discretization and Equal width binning

#### 3.1.1 2 bins

##### Cross Validation 5 folds

```
==== Run information ====
Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B2-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  30
              age_1
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              T3_measured
              T3_1
              TT4_measured
              TT4_1
              T4U_measured
              T4U_1
              FTI_measured
              FTI_1
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   5-fold cross-validation
==== Classifier model (full training set) ====
PART decision list
-----
referral_source = other AND
T3_measured = t AND
sick = f AND
T4U_measured = t AND
T4U_1 = '(-inf-1.285]' AND
tumor = f AND
on_antithyroid_medication = f AND
thyroid_surgery = f AND
psych = f AND
TT4_1 = '(-inf-216]' AND
FTI_1 = '(-inf-198.5]': negative (1108.0/27.0)
```

```

referral_source = other: negative (1093.0/6.0)

referral_source = SVHC AND
query_hypothyroid = f: negative (376.0/6.0)

referral_source = SVI AND
thyroid_surgery = f AND
lithium = f AND
TT4_measured = t AND
psych = f AND
T3_measured = t AND
TT4_1 = '(-inf-216]' AND
query_hypothyroid = f AND
T4U_1 = '(-inf-1.285]' AND
sick = f AND
TSH_measured = t AND
FTI_measured = t AND
FTI_1 = '(-inf-198.5]' AND
query_hyperthyroid = f AND
query_on_thyroxine = f AND
goitre = f AND
sex = F: negative (436.41/73.28)

referral_source = SVI AND
thyroid_surgery = f AND
TT4_measured = t AND
T3_measured = t AND
T4U_1 = '(-inf-1.285]' AND
tumor = f AND
psych = f: negative (503.39/104.43)

psych = f: negative (238.2/8.29)

sick = f: negative (15.0/4.0)

: sick (2.0)

Number of Rules : 8

Time taken to build model: 0.11 seconds

==== Stratified cross-validation ===
==== Summary ===

    Correctly Classified Instances      3536           93.7434 %
    Incorrectly Classified Instances   236            6.2566 %
    Kappa statistic                   0.0121
    Mean absolute error              0.1028
    Root mean squared error          0.2291
    Relative absolute error          89.2466 %
    Root relative squared error     95.5371 %
    Total Number of Instances       3772

==== Detailed Accuracy By Class ===

          TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
          0.998     0.991     0.939     0.998     0.968     0.033     0.799     0.980     negative
          0.009     0.002     0.222     0.009     0.017     0.033     0.799     0.183     sick
Weighted Avg.   0.937     0.931     0.895     0.937     0.909     0.033     0.799     0.932

==== Confusion Matrix ===

      a      b  <-- classified as
3534    7 |  a = negative
 229    2 |  b = sick

```

## Cross Validation 10 folds

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B2-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  30
             age_1
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery

```

```

I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH_1
T3_measured
T3_1
TT4_measured
TT4_1
T4U_measured
T4U_1
FTI_measured
FTI_1
TBG_measured
TBG
referral_source
Class
Test mode: 10-fold cross-validation

==== Classifier model (full training set) ====

PART decision list
-----
referral_source = other AND
T3_measured = t AND
sick = f AND
T4U_measured = t AND
T4U_1 = '(-inf-1.285]' AND
tumor = f AND
on_antithyroid_medication = f AND
thyroid_surgery = f AND
psych = f AND
TT4_1 = '(-inf-216]' AND
FTI_1 = '(-inf-198.5]': negative (1108.0/27.0)

referral_source = other: negative (1093.0/6.0)

referral_source = SVHC AND
query_hypothyroid = f: negative (376.0/6.0)

referral_source = SVI AND
thyroid_surgery = f AND
lithium = f AND
TT4_measured = t AND
psych = f AND
T3_measured = t AND
TT4_1 = '(-inf-216]' AND
query_hypothyroid = f AND
T4U_1 = '(-inf-1.285]' AND
sick = f AND
TSH_measured = t AND
FTI_measured = t AND
FTI_1 = '(-inf-198.5]' AND
query_hyperthyroid = f AND
query_on_thyroxine = f AND
goitre = f AND
sex = F: negative (436.41/73.28)

referral_source = SVI AND
thyroid_surgery = f AND
TT4_measured = t AND
T3_measured = t AND
T4U_1 = '(-inf-1.285]' AND
tumor = f AND
psych = f: negative (503.39/104.43)

psych = f: negative (238.2/8.29)

sick = f: negative (15.0/4.0)

: sick (2.0)

Number of Rules : 8

Time taken to build model: 0.1 seconds

==== Stratified cross-validation ====
==== Summary ====
Correctly Classified Instances          3540           93.8494 %

```

```

Incorrectly Classified Instances          232           6.1506 %
Kappa statistic                      0.007
Mean absolute error                  0.1031
Root mean squared error              0.2286
Relative absolute error              89.4994 %
Root relative squared error         95.3291 %
Total Number of Instances            3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.999     0.996     0.939     0.999     0.968     0.032     0.798     0.981    negative
      0.004     0.001     0.333     0.004     0.009     0.032     0.798     0.177    sick
Weighted Avg.      0.938     0.935     0.902     0.938     0.909     0.032     0.798     0.931

==== Confusion Matrix ====

      a     b  <-- classified as
3539  2 |    a = negative
 230  1 |    b = sick

```

## Cross Validation 15 folds

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B2-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  30
             age_1
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             I131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary
             psych
             TSH_measured
             TSH_1
             T3_measured
             T3_1
             TT4_measured
             TT4_1
             T4U_measured
             T4U_1
             FTI_measured
             FTI_1
             TBG_measured
             TBG
             referral_source
             Class
Test mode:   15-fold cross-validation

==== Classifier model (full training set) ===

PART decision list
-----
referral_source = other AND
T3_measured = t AND
sick = f AND
T4U_measured = t AND
T4U_1 = '(-inf-1.285]' AND
tumor = f AND
on_antithyroid_medication = f AND
thyroid_surgery = f AND
psych = f AND
TT4_1 = '(-inf-216]' AND
FTI_1 = '(-inf-198.5]': negative (1108.0/27.0)

referral_source = other: negative (1093.0/6.0)

referral_source = SVHC AND
query_hypothyroid = f: negative (376.0/6.0)

referral_source = SVI AND

```

```

thyroid_surgery = f AND
lithium = f AND
TT4_measured = t AND
psych = f AND
T3_measured = t AND
TT4_1 = '(-inf-216]' AND
query_hypothyroid = f AND
T4U_1 = '(-inf-1.285]' AND
sick = f AND
TSH_measured = t AND
FTI_measured = t AND
FTI_1 = '(-inf-198.5]' AND
query_hyperthyroid = f AND
query_on_thyroxine = f AND
goitre = f AND
sex = F: negative (436.41/73.28)

referral_source = SVI AND
thyroid_surgery = f AND
TT4_measured = t AND
T3_measured = t AND
T4U_1 = '(-inf-1.285]' AND
tumor = f AND
psych = f: negative (503.39/104.43)

psych = f: negative (238.2/8.29)

sick = f: negative (15.0/4.0)

: sick (2.0)

Number of Rules : 8

Time taken to build model: 0.1 seconds

==== Stratified cross-validation ====
==== Summary ===

    Correctly Classified Instances      3539          93.8229 %
    Incorrectly Classified Instances   233           6.1771 %
    Kappa statistic                   -0.0011
    Mean absolute error              0.1034
    Root mean squared error         0.2292
    Relative absolute error          89.733  %
    Root relative squared error     95.598  %
    Total Number of Instances       3772

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.999     1.000     0.939      0.999     0.968     -0.006    0.793     0.981     negative
      0.000     0.001     0.000      0.000     0.000     -0.006    0.793     0.171     sick
Weighted Avg.      0.938     0.939     0.881      0.938     0.909     -0.006    0.793     0.931

==== Confusion Matrix ===

      a     b  <-- classified as
3539   2 |   a = negative
  231   0 |   b = sick

```

## Percentage Split 50%

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B2-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  30
             age_1
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             i131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary

```

```

psych
TSH_measured
TSH_1
T3_measured
T3_1
TT4_measured
TT4_1
T4U_measured
T4U_1
FTI_measured
FTI_1
TBG_measured
TBG
referral_source
Class
Test mode: split 50.0% train, remainder test

==== Classifier model (full training set) ===

PART decision list
-----
referral_source = other AND
T3_measured = t AND
sick = f AND
T4U_measured = t AND
T4U_1 = '(-inf-1.285)' AND
tumor = f AND
on_antithyroid_medication = f AND
thyroid_surgery = f AND
psych = f AND
TT4_1 = '(-inf-216)' AND
FTI_1 = '(-inf-198.5)': negative (1108.0/27.0)

referral_source = other: negative (1093.0/6.0)

referral_source = SVHC AND
query_hypothyroid = f: negative (376.0/6.0)

referral_source = SVI AND
thyroid_surgery = f AND
lithium = f AND
TT4_measured = t AND
psych = f AND
T3_measured = t AND
TT4_1 = '(-inf-216)' AND
query_hypothyroid = f AND
T4U_1 = '(-inf-1.285)' AND
sick = f AND
TSH_measured = t AND
FTI_measured = t AND
FTI_1 = '(-inf-198.5)' AND
query_hyperthyroid = f AND
query_on_thyroxine = f AND
goitre = f AND
sex = F: negative (436.41/73.28)

referral_source = SVI AND
thyroid_surgery = f AND
TT4_measured = t AND
T3_measured = t AND
T4U_1 = '(-inf-1.285)' AND
tumor = f AND
psych = f: negative (503.39/104.43)

psych = f: negative (238.2/8.29)

sick = f: negative (15.0/4.0)

: sick (2.0)

Number of Rules : 8

Time taken to build model: 0.1 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

==== Summary ===

Correctly Classified Instances      1768          93.7434 %
Incorrectly Classified Instances   118           6.2566 %
Kappa statistic                   -0.0011
Mean absolute error               0.1038
Root mean squared error           0.2307

```

```

Relative absolute error          89.9291 %
Root relative squared error    95.6233 %
Total Number of Instances      1886

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.999     1.000     0.938     0.999     0.968     -0.006    0.783     0.974     negative
      0.000     0.001     0.000     0.000     0.000     -0.006    0.783     0.182     sick
Weighted Avg.      0.937     0.938     0.880     0.937     0.908     -0.006    0.783     0.925

==== Confusion Matrix ====

      a      b  <-- classified as
1768    1 |  a = negative
 117    0 |  b = sick

```

## Percentage Split 60%

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B2-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  30
              age_1
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              T3_measured
              T3_1
              TT4_measured
              TT4_1
              T4U_measured
              T4U_1
              FTI_measured
              FTI_1
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 60.0% train, remainder test

==== Classifier model (full training set) ===

PART decision list
-----
referral_source = other AND
T3_measured = t AND
sick = f AND
T4U_measured = t AND
T4U_1 = '(-inf-1.285]' AND
tumor = f AND
on_antithyroid_medication = f AND
thyroid_surgery = f AND
psych = f AND
TT4_1 = '(-inf-216]' AND
FTI_1 = '(-inf-198.5]': negative (1108.0/27.0)

referral_source = other: negative (1093.0/6.0)

referral_source = SVHC AND
query_hypothyroid = f: negative (376.0/6.0)

referral_source = SVI AND
thyroid_surgery = f AND
lithium = f AND
TT4_measured = t AND
psych = f AND

```

```

T3_measured = t AND
TT4_1 = '(-inf-216]' AND
query_hypothyroid = f AND
T4U_1 = '(-inf-1.285]' AND
sick = f AND
TSH_measured = t AND
FTI_measured = t AND
FTI_1 = '(-inf-198.5]' AND
query_hyperthyroid = f AND
query_on_thyroxine = f AND
goitre = f AND
sex = F: negative (436.41/73.28)

referral_source = SVI AND
thyroid_surgery = f AND
TT4_measured = t AND
T3_measured = t AND
T4U_1 = '(-inf-1.285]' AND
tumor = f AND
psych = f: negative (503.39/104.43)

psych = f: negative (238.2/8.29)

sick = f: negative (15.0/4.0)

: sick (2.0)

Number of Rules : 8

Time taken to build model: 0.11 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

==== Summary ===

Correctly Classified Instances      1420           94.1021 %
Incorrectly Classified Instances   89            5.8979 %
Kappa statistic                   0.0561
Mean absolute error               0.1011
Root mean squared error          0.2249
Relative absolute error          88.1546 %
Root relative squared error     95.4659 %
Total Number of Instances        1509

==== Detailed Accuracy By Class ===

          TP Rate  FP Rate  Precision  Recall  F-Measure  MCC    ROC Area  PRC Area  Class
          0.998    0.966    0.943    0.998    0.970    0.118    0.779    0.975    negative
          0.034    0.002    0.500    0.034    0.063    0.118    0.779    0.183    sick
Weighted Avg.    0.941    0.909    0.917    0.941    0.916    0.118    0.779    0.928

==== Confusion Matrix ===

      a      b  <-- classified as
1417    3 |      a = negative
      86    3 |      b = sick

```

## Percentage Split 66%

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B2-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  30
              age_1
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych

```

```

TSH_measured
TSH_1
T3_measured
T3_1
TT4_measured
TT4_1
T4U_measured
T4U_1
FTI_measured
FTI_1
TBG_measured
TBG
referral_source
Class
Test mode: split 66.0% train, remainder test

==== Classifier model (full training set) ====

PART decision list
-----
referral_source = other AND
T3_measured = t AND
sick = f AND
T4U_measured = t AND
T4U_1 = '(-inf-1.285]' AND
tumor = f AND
on_antithyroid_medication = f AND
thyroid_surgery = f AND
psych = f AND
TT4_1 = '(-inf-216]' AND
FTI_1 = '(-inf-198.5]': negative (1108.0/27.0)

referral_source = other: negative (1093.0/6.0)

referral_source = SVHC AND
query_hypothyroid = f: negative (376.0/6.0)

referral_source = SVI AND
thyroid_surgery = f AND
lithium = f AND
TT4_measured = t AND
psych = f AND
T3_measured = t AND
TT4_1 = '(-inf-216]' AND
query_hypothyroid = f AND
T4U_1 = '(-inf-1.285]' AND
sick = f AND
TSH_measured = t AND
FTI_measured = t AND
FTI_1 = '(-inf-198.5]' AND
query_hyperthyroid = f AND
query_on_thyroxine = f AND
goitre = f AND
sex = F: negative (436.41/73.28)

referral_source = SVI AND
thyroid_surgery = f AND
TT4_measured = t AND
T3_measured = t AND
T4U_1 = '(-inf-1.285]' AND
tumor = f AND
psych = f: negative (503.39/104.43)

psych = f: negative (238.2/8.29)

sick = f: negative (15.0/4.0)

: sick (2.0)

Number of Rules : 8

Time taken to build model: 0.1 seconds

==== Evaluation on test split ====

Time taken to test model on test split: 0 seconds

==== Summary ====

Correctly Classified Instances      1196          93.2917 %
Incorrectly Classified Instances    86           6.7083 %
Kappa statistic                   -0.0152
Mean absolute error                0.1042
Root mean squared error            0.2337
Relative absolute error             91.3181 %

```

```

Root relative squared error          99.5579 %
Total Number of Instances          1282

==== Detailed Accuracy By Class ====

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
          0.991    1.000     0.941     0.991     0.965    -0.023    0.756     0.973  negative
          0.000    0.009     0.000     0.000     0.000    -0.023    0.756     0.139    sick
Weighted Avg.     0.933     0.942     0.886     0.933     0.909    -0.023    0.756     0.924

==== Confusion Matrix ====

     a     b  <-- classified as
1196  11 |     a = negative
    75   0 |     b = sick

```

## Leave one out fold

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B2-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  30
              age_1
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              t131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              T3_measured
              T3_1
              TT4_measured
              TT4_1
              T4U_measured
              T4U_1
              FTI_measured
              FTI_1
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   3772-fold cross-validation

==== Classifier model (full training set) ===

PART decision list
-----
referral_source = other AND
T3_measured = t AND
sick = f AND
T4U_measured = t AND
T4U_1 = '(-inf-1.285]' AND
tumor = f AND
on_antithyroid_medication = f AND
thyroid_surgery = f AND
psych = f AND
TT4_1 = '(-inf-216]' AND
FTI_1 = '(-inf-198.5]': negative (1108.0/27.0)

referral_source = other: negative (1093.0/6.0)

referral_source = SVHC AND
query_hypothyroid = f: negative (376.0/6.0)

referral_source = SVI AND
thyroid_surgery = f AND
lithium = f AND
TT4_measured = t AND
psych = f AND
T3_measured = t AND

```

```

TT4_1 = '(-inf-216)' AND
query_hypothyroid = f AND
T4U_1 = '(-inf-1.285)' AND
sick = f AND
TSH_measured = t AND
FTI_measured = t AND
FTI_1 = '(-inf-198.5)' AND
query_hyperthyroid = f AND
query_on_thyroxine = f AND
goitre = f AND
sex = F: negative (436.41/73.28)

referral_source = SVI AND
thyroid_surgery = f AND
TT4_measured = t AND
T3_measured = t AND
T4U_1 = '(-inf-1.285)' AND
tumor = f AND
psych = f: negative (503.39/104.43)

psych = f: negative (238.2/8.29)

sick = f: negative (15.0/4.0)

: sick (2.0)

Number of Rules : 8

Time taken to build model: 0.12 seconds

==== Stratified cross-validation ====
==== Summary ====

    Correctly Classified Instances      3540           93.8494 %
    Incorrectly Classified Instances   232            6.1506 %
    Kappa statistic                   0.007
    Mean absolute error              0.103
    Root mean squared error          0.2281
    Relative absolute error          89.4271 %
    Root relative squared error     95.1223 %
    Total Number of Instances        3772

==== Detailed Accuracy By Class ====

          TP Rate   FP Rate   Precision   Recall   F-Measure   MCC     ROC Area   PRC Area   Class
          0.999     0.996     0.939     0.999     0.968     0.032     0.755     0.980   negative
          0.004     0.001     0.333     0.004     0.009     0.032     0.755     0.143   sick
Weighted Avg.   0.938     0.935     0.902     0.938     0.909     0.032     0.755     0.929

==== Confusion Matrix ====

      a      b  <-- classified as
3539  2 |  a = negative
 230  1 |  b = sick

```

### 3.1.2 5 bins

#### Cross Validation 5 folds

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B5-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  48
             age_1
             age_2
             age_3
             age_4
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             I131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor

```

```

hypopituitary
psych
TSH_measured
TSH_1
TSH_2
TSH_3
TSH_4
T3_measured
T3_1
T3_2
T3_3
T3_4
TT4_measured
TT4_1
TT4_2
TT4_3
TT4_4
T4U_measured
T4U_1
T4U_2
T4U_3
T4U_4
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
TBG_measured
TBG
referral_source
Class
Test mode: 5-fold cross-validation

==== Classifier model (full training set) ====

PART decision list
-----
referral_source = other AND
age_1 = '(-inf-91.8]' AND
FTI_2 = '(-inf-159.2]' AND
query_hypothyroid = f: negative (1878.08/16.0)

T3_1 = '(2.16-inf)': negative (656.23/7.04)

T3_measured = t AND
referral_source = SVHC AND
query_hypothyroid = f: negative (194.0/6.0)

T3_measured = t AND
thyroid_surgery = f AND
psych = f AND
TT4_1 = '(87.6-inf)' AND
TSH_measured = t AND
TT4_2 = '(-inf-173.2)': negative (585.16/97.0)

T3_measured = f: negative (126.02/1.96)

thyroid_surgery = f AND
T4U_measured = t AND
T4U_2 = '(1.078-inf)': negative (32.0/2.0)

thyroid_surgery = f AND
T4U_measured = f: negative (16.84/1.0)

thyroid_surgery = f AND
query_hyperthyroid = f AND
TSH_1 = '(-inf-106.004]' AND
on_thyroxine = f AND
sick = t AND
referral_source = SVI: sick (16.0/4.0)

thyroid_surgery = f AND
query_hyperthyroid = f AND
TSH_1 = '(-inf-106.004]' AND
on_thyroxine = f AND
referral_source = SVI AND
T4U_1 = '(0.664-inf)' AND
sex = F: negative (101.09/28.1)

thyroid_surgery = f AND
psych = f AND
query_hyperthyroid = f AND
TSH_1 = '(-inf-106.004]' AND
on_thyroxine = f AND
TT4_2 = '(-inf-173.2]' AND
referral_source = SVI AND

```

```

query_on_thyroxine = f AND
FTI_1 = '(80.6-inf)' AND
T4U_1 = '(0.664-inf)': negative (56.46/16.45)

thyroid_surgery = f AND
query_hyperthyroid = f AND
on_thyroxine = f AND
TSH_1 = '(-inf-106.004]' AND
referral_source = SVI AND
psych = f AND
sex = M: sick (35.05/12.72)

psych = f AND
thyroid_surgery = f AND
sex = M: negative (14.69/1.03)

sex = F AND
thyroid_surgery = f AND
query_hyperthyroid = f AND
FTI_1 = '(-inf-80.6]': negative (13.8/3.09)

sex = F AND
thyroid_surgery = t: negative (9.32)

query_hyperthyroid = f AND
sex = F AND
on_thyroxine = t: negative (6.0/1.0)

query_hyperthyroid = f AND
sex = F AND
referral_source = other AND
query_hypothyroid = f: sick (5.0)

sex = F AND
query_hyperthyroid = f AND
TT4_1 = '(-inf-87.6]' AND
psych = f: negative (14.22/6.0)

psych = f: negative (6.03/0.0)

: sick (6.0/1.0)

Number of Rules : 19

Time taken to build model: 0.17 seconds

==== Stratified cross-validation ====
==== Summary ====

    Correctly Classified Instances      3540           93.8494 %
    Incorrectly Classified Instances   232            6.1506 %
    Kappa statistic                   0.2137
    Mean absolute error              0.094
    Root mean squared error          0.2237
    Relative absolute error          81.5667 %
    Root relative squared error     93.3037 %
    Total Number of Instances        3772

==== Detailed Accuracy By Class ====

          TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
          0.990     0.844     0.947     0.990     0.968     0.253     0.831     0.980     negative
          0.156     0.010     0.493     0.156     0.237     0.253     0.831     0.271     sick
Weighted Avg.   0.938     0.793     0.919     0.938     0.923     0.253     0.831     0.936

==== Confusion Matrix ====

      a     b  <-- classified as
3504   37 |   a = negative
  95   36 |   b = sick

```

## Cross Validation 10 folds

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B5-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  48
             age_1
             age_2
             age_3
             age_4

```

```

sex
on_thyroxine
query_on_thyroxine
on_antithyroid_medication
sick
pregnant
thyroid_surgery
I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH_1
TSH_2
TSH_3
TSH_4
T3_measured
T3_1
T3_2
T3_3
T3_4
TT4_measured
TT4_1
TT4_2
TT4_3
TT4_4
T4U_measured
T4U_1
T4U_2
T4U_3
T4U_4
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
TBG_measured
TBG
referral_source
Class
Test mode: 10-fold cross-validation

==== Classifier model (full training set) ====

PART decision list
-----

referral_source = other AND
age_1 = '(-inf-91.8]' AND
FTI_2 = '(-inf-159.2]' AND
query_hypothyroid = f: negative (1878.08/16.0)

T3_1 = '(2.16-inf)': negative (656.23/7.04)

T3_measured = t AND
referral_source = SVHC AND
query_hypothyroid = f: negative (194.0/6.0)

T3_measured = t AND
thyroid_surgery = f AND
psych = f AND
TT4_1 = '(87.6-inf)' AND
TSH_measured = t AND
TT4_2 = '(-inf-173.2)': negative (585.16/97.0)

T3_measured = f: negative (126.02/1.96)

thyroid_surgery = f AND
T4U_measured = t AND
T4U_2 = '(1.078-inf)': negative (32.0/2.0)

thyroid_surgery = f AND
T4U_measured = f: negative (16.84/1.0)

thyroid_surgery = f AND
query_hyperthyroid = f AND
TSH_1 = '(-inf-106.004]' AND
on_thyroxine = f AND
sick = t AND
referral_source = SVI: sick (16.0/4.0)

thyroid_surgery = f AND
query_hyperthyroid = f AND

```

```

TSH_1 = '(-inf-106.004]' AND
on_thyroxine = f AND
referral_source = SVI AND
T4U_1 = '(0.664-inf)' AND
sex = F: negative (101.09/28.1)

thyroid_surgery = f AND
psych = f AND
query_hyperthyroid = f AND
TSH_1 = '(-inf-106.004]' AND
on_thyroxine = f AND
TT4_2 = '(-inf-173.2]' AND
referral_source = SVI AND
query_on_thyroxine = f AND
FTI_1 = '(80.6-inf)' AND
T4U_1 = '(0.664-inf)': negative (56.46/16.45)

thyroid_surgery = f AND
query_hyperthyroid = f AND
on_thyroxine = f AND
TSH_1 = '(-inf-106.004]' AND
referral_source = SVI AND
psych = f AND
sex = M: sick (35.05/12.72)

psych = f AND
thyroid_surgery = f AND
sex = M: negative (14.69/1.03)

sex = F AND
thyroid_surgery = f AND
query_hyperthyroid = f AND
FTI_1 = '(-inf-80.6]': negative (13.8/3.09)

sex = F AND
thyroid_surgery = t: negative (9.32)

query_hyperthyroid = f AND
sex = F AND
on_thyroxine = t: negative (6.0/1.0)

query_hyperthyroid = f AND
sex = F AND
referral_source = other AND
query_hypothyroid = f: sick (5.0)

sex = F AND
query_hyperthyroid = f AND
TT4_1 = '(-inf-87.6]' AND
psych = f: negative (14.22/6.0)

psych = f: negative (6.03/0.0)

: sick (6.0/1.0)

Number of Rules : 19

Time taken to build model: 0.14 seconds

==== Stratified cross-validation ===
==== Summary ===

Correctly Classified Instances      3552           94.1676 %
Incorrectly Classified Instances   220            5.8324 %
Kappa statistic                   0.2408
Mean absolute error               0.0917
Root mean squared error          0.2196
Relative absolute error           79.6318 %
Root relative squared error      91.5881 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall    F-Measure   MCC      ROC Area   PRC Area   Class
      0.992     0.831     0.948     0.992     0.970      0.292     0.839     0.982   negative
      0.169     0.008     0.582     0.169     0.262      0.292     0.839     0.302     sick
Weighted Avg.      0.942     0.781     0.926     0.942     0.926      0.292     0.839     0.941

==== Confusion Matrix ====

      a      b  <-- classified as
3513   28 |      a = negative
      192   39 |      b = sick

```

## Cross Validation 15 folds

```
==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B5-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  48
              age_1
              age_2
              age_3
              age_4
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              t131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              TSH_2
              TSH_3
              TSH_4
              T3_measured
              T3_1
              T3_2
              T3_3
              T3_4
              TT4_measured
              TT4_1
              TT4_2
              TT4_3
              TT4_4
              T4U_measured
              T4U_1
              T4U_2
              T4U_3
              T4U_4
              FTI_measured
              FTI_1
              FTI_2
              FTI_3
              FTI_4
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   15-fold cross-validation

==== Classifier model (full training set) ===

PART decision list
-----

referral_source = other AND
age_1 = '(-inf-91.8]' AND
FTI_2 = '(-inf-159.2]' AND
query_hypothyroid = f: negative (1878.08/16.0)

T3_1 = '(2.16-inf)': negative (656.23/7.04)

T3_measured = t AND
referral_source = SVHC AND
query_hypothyroid = f: negative (194.0/6.0)

T3_measured = t AND
thyroid_surgery = f AND
psych = f AND
TT4_1 = '(87.6-inf)' AND
TSH_measured = t AND
TT4_2 = '(-inf-173.2)': negative (585.16/97.0)

T3_measured = f: negative (126.02/1.96)

thyroid_surgery = f AND
T4U_measured = t AND
T4U_2 = '(1.078-inf)': negative (32.0/2.0)
```

```

thyroid_surgery = f AND
T4U_measured = f: negative (16.84/1.0)

thyroid_surgery = f AND
query_hyperthyroid = f AND
TSH_1 = '(-inf-106.004]' AND
on_thyroxine = f AND
sick = t AND
referral_source = SVI: sick (16.0/4.0)

thyroid_surgery = f AND
query_hyperthyroid = f AND
TSH_1 = '(-inf-106.004]' AND
on_thyroxine = f AND
referral_source = SVI AND
T4U_1 = '(0.664-inf)' AND
sex = F: negative (101.09/28.1)

thyroid_surgery = f AND
psych = f AND
query_hyperthyroid = f AND
TSH_1 = '(-inf-106.004]' AND
on_thyroxine = f AND
TT4_2 = '(-inf-173.2]' AND
referral_source = SVI AND
query_on_thyroxine = f AND
FTI_1 = '(80.6-inf)' AND
T4U_1 = '(0.664-inf)': negative (56.46/16.45)

thyroid_surgery = f AND
query_hyperthyroid = f AND
on_thyroxine = f AND
TSH_1 = '(-inf-106.004]' AND
referral_source = SVI AND
psych = f AND
sex = M: sick (35.05/12.72)

psych = f AND
thyroid_surgery = f AND
sex = M: negative (14.69/1.03)

sex = F AND
thyroid_surgery = f AND
query_hyperthyroid = f AND
FTI_1 = '(-inf-80.6]': negative (13.8/3.09)

sex = F AND
thyroid_surgery = t: negative (9.32)

query_hyperthyroid = f AND
sex = F AND
on_thyroxine = t: negative (6.0/1.0)

query_hyperthyroid = f AND
sex = F AND
referral_source = other AND
query_hypothyroid = f: sick (5.0)

sex = F AND
query_hyperthyroid = f AND
TT4_1 = '(-inf-87.6]' AND
psych = f: negative (14.22/6.0)

psych = f: negative (6.03/0.0)

: sick (6.0/1.0)

Number of Rules : 19

Time taken to build model: 0.2 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3550          94.1145 %
Incorrectly Classified Instances   222           5.8855 %
Kappa statistic                   0.2245
Mean absolute error               0.0918
Root mean squared error          0.2202
Relative absolute error           79.6584 %
Root relative squared error      91.8296 %
Total Number of Instances         3772

==== Detailed Accuracy By Class ====

```

TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
---------	---------	-----------	--------	-----------	-----	----------	----------	-------

```

0.992    0.844    0.947    0.992    0.969    0.277    0.827    0.977    negative
0.156    0.008    0.571    0.156    0.245    0.277    0.827    0.296    sick
Weighted Avg. 0.941    0.793    0.924    0.941    0.925    0.277    0.827    0.935

==== Confusion Matrix ====
a      b    <-- classified as
3514   27 |     a = negative
195    36 |     b = sick

```

## Percentage Split 50%

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B5-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  48
              age_1
              age_2
              age_3
              age_4
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              TSH_2
              TSH_3
              TSH_4
              T3_measured
              T3_1
              T3_2
              T3_3
              T3_4
              TT4_measured
              TT4_1
              TT4_2
              TT4_3
              TT4_4
              T4U_measured
              T4U_1
              T4U_2
              T4U_3
              T4U_4
              FTI_measured
              FTI_1
              FTI_2
              FTI_3
              FTI_4
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 50.0% train, remainder test

==== Classifier model (full training set) ===

PART decision list
-----
referral_source = other AND
age_1 = '(-inf-91.8]' AND
FTI_2 = '(-inf-159.2]' AND
query_hypothyroid = f: negative (1878.08/16.0)

T3_1 = '(2.16-inf)': negative (656.23/7.04)

T3_measured = t AND
referral_source = SVHC AND
query_hypothyroid = f: negative (194.0/6.0)

```

```

T3_measured = t AND
thyroid_surgery = f AND
psych = f AND
TT4_1 = '(87.6-inf)' AND
TSH_measured = t AND
TT4_2 = '(-inf-173.2)': negative (585.16/97.0)

T3_measured = f: negative (126.02/1.96)

thyroid_surgery = f AND
T4U_measured = t AND
T4U_2 = '(1.078-inf)': negative (32.0/2.0)

thyroid_surgery = f AND
T4U_measured = f: negative (16.84/1.0)

thyroid_surgery = f AND
query_hyperthyroid = f AND
TSH_1 = '(-inf-106.004]' AND
on_thyroxine = f AND
sick = t AND
referral_source = SVI: sick (16.0/4.0)

thyroid_surgery = f AND
query_hyperthyroid = f AND
TSH_1 = '(-inf-106.004]' AND
on_thyroxine = f AND
referral_source = SVI AND
T4U_1 = '(0.664-inf)' AND
sex = F: negative (101.09/28.1)

thyroid_surgery = f AND
psych = f AND
query_hyperthyroid = f AND
TSH_1 = '(-inf-106.004]' AND
on_thyroxine = f AND
TT4_2 = '(-inf-173.2)' AND
referral_source = SVI AND
query_on_thyroxine = f AND
FTI_1 = '(80.6-inf)' AND
T4U_1 = '(0.664-inf)': negative (56.46/16.45)

thyroid_surgery = f AND
query_hyperthyroid = f AND
on_thyroxine = f AND
TSH_1 = '(-inf-106.004]' AND
referral_source = SVI AND
psych = f AND
sex = M: sick (35.05/12.72)

psych = f AND
thyroid_surgery = f AND
sex = M: negative (14.69/1.03)

sex = F AND
thyroid_surgery = f AND
query_hyperthyroid = f AND
FTI_1 = '(-inf-80.6)': negative (13.8/3.09)

sex = F AND
thyroid_surgery = t: negative (9.32)

query_hyperthyroid = f AND
sex = F AND
on_thyroxine = t: negative (6.0/1.0)

query_hyperthyroid = f AND
sex = F AND
referral_source = other AND
query_hypothyroid = f: sick (5.0)

sex = F AND
query_hyperthyroid = f AND
TT4_1 = '(-inf-87.6)' AND
psych = f: negative (14.22/6.0)

psych = f: negative (6.03/0.0)

: sick (6.0/1.0)

Number of Rules : 19

Time taken to build model: 0.16 seconds
== Evaluation on test split ==

```

```

Time taken to test model on test split: 0 seconds

==== Summary ====
Correctly Classified Instances      1769          93.7964 %
Incorrectly Classified Instances    117           6.2036 %
Kappa statistic                      0
Mean absolute error                  0.0956
Root mean squared error              0.2222
Relative absolute error              82.8183 %
Root relative squared error         92.0962 %
Total Number of Instances           1886

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC   ROC Area   PRC Area   Class
      1.000     1.000     0.938      1.000     0.968     ?     0.846      0.983   negative
      0.000     0.000     ?          0.000     ?          ?     0.846      0.259   sick
Weighted Avg.   0.938     0.938     ?          0.938     ?          ?     0.846      0.938

==== Confusion Matrix ====

      a      b  <-- classified as
1769  0 |  a = negative
  117  0 |  b = sick

```

## Percentage Split 60%

```

==== Run information ====
Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B5-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  48
             age_1
             age_2
             age_3
             age_4
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             I131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary
             psych
             TSH_measured
             TSH_1
             TSH_2
             TSH_3
             TSH_4
             T3_measured
             T3_1
             T3_2
             T3_3
             T3_4
             TT4_measured
             TT4_1
             TT4_2
             TT4_3
             TT4_4
             T4U_measured
             T4U_1
             T4U_2
             T4U_3
             T4U_4
             FTI_measured
             FTI_1
             FTI_2
             FTI_3
             FTI_4
             TBG_measured
             TBG
             referral_source
             Class
Test mode:   split 60.0% train, remainder test

```

```

==== Classifier model (full training set) ====

PART decision list
-----
referral_source = other AND
age_1 = '(-inf-91.8]' AND
FTI_2 = '(-inf-159.2]' AND
query_hypothyroid = f: negative (1878.08/16.0)

T3_1 = '(2.16-inf)': negative (656.23/7.04)

T3_measured = t AND
referral_source = SVHC AND
query_hypothyroid = f: negative (194.0/6.0)

T3_measured = t AND
thyroid_surgery = f AND
psych = f AND
TT4_1 = '(87.6-inf)' AND
TSH_measured = t AND
TT4_2 = '(-inf-173.2)': negative (585.16/97.0)

T3_measured = f: negative (126.02/1.96)

thyroid_surgery = f AND
T4U_measured = t AND
T4U_2 = '(1.078-inf)': negative (32.0/2.0)

thyroid_surgery = f AND
T4U_measured = f: negative (16.84/1.0)

thyroid_surgery = f AND
query_hyperthyroid = f AND
TSH_1 = '(-inf-106.004]' AND
on_thyroxine = f AND
sick = t AND
referral_source = SVI: sick (16.0/4.0)

thyroid_surgery = f AND
query_hyperthyroid = f AND
TSH_1 = '(-inf-106.004]' AND
on_thyroxine = f AND
referral_source = SVI AND
T4U_1 = '(0.664-inf)' AND
sex = F: negative (101.09/28.1)

thyroid_surgery = f AND
psych = f AND
query_hyperthyroid = f AND
TSH_1 = '(-inf-106.004]' AND
on_thyroxine = f AND
TT4_2 = '(-inf-173.2]' AND
referral_source = SVI AND
query_on_thyroxine = f AND
FTI_1 = '(80.6-inf)' AND
T4U_1 = '(0.664-inf)': negative (56.46/16.45)

thyroid_surgery = f AND
query_hyperthyroid = f AND
on_thyroxine = f AND
TSH_1 = '(-inf-106.004]' AND
referral_source = SVI AND
psych = f AND
sex = M: sick (35.05/12.72)

psych = f AND
thyroid_surgery = f AND
sex = M: negative (14.69/1.03)

sex = F AND
thyroid_surgery = f AND
query_hyperthyroid = f AND
FTI_1 = '(-inf-80.6)': negative (13.8/3.09)

sex = F AND
thyroid_surgery = t: negative (9.32)

query_hyperthyroid = f AND
sex = F AND
on_thyroxine = t: negative (6.0/1.0)

query_hyperthyroid = f AND
sex = F AND
referral_source = other AND
query_hypothyroid = f: sick (5.0)

```

```

sex = F AND
query_hyperthyroid = f AND
TT4_1 = '(-inf-87.6]' AND
psych = f: negative (14.22/6.0)

psych = f: negative (6.03/0.0)

: sick (6.0/1.0)

Number of Rules : 19

Time taken to build model: 0.15 seconds

==== Evaluation on test split ====

Time taken to test model on test split: 0 seconds

==== Summary ====

Correctly Classified Instances      1419          94.0358 %
Incorrectly Classified Instances   90           5.9642 %
Kappa statistic                   0.149
Mean absolute error               0.0893
Root mean squared error          0.2196
Relative absolute error          77.8782 %
Root relative squared error     93.2002 %
Total Number of Instances        1509

==== Detailed Accuracy By Class ====

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC    ROC Area  PRC Area  Class
      0.993    0.899    0.946     0.993    0.969     0.199    0.850    0.982    negative
      0.101    0.007    0.474     0.101    0.167     0.199    0.850    0.265    sick
Weighted Avg.      0.940    0.846    0.918     0.940    0.922     0.199    0.850    0.940

==== Confusion Matrix ====

      a      b  <-- classified as
1410  10 |      a = negative
      80   9 |      b = sick

```

## Percentage Split 66%

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B5-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  48
             age_1
             age_2
             age_3
             age_4
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             I131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary
             psych
             TSH_measured
             TSH_1
             TSH_2
             TSH_3
             TSH_4
             T3_measured
             T3_1
             T3_2
             T3_3
             T3_4
             TT4_measured
             TT4_1
             TT4_2
             TT4_3
             TT4_4

```

```

T4U_measured
T4U_1
T4U_2
T4U_3
T4U_4
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
TBG_measured
TBG
referral_source
Class
Test mode: split 66.0% train, remainder test

==== Classifier model (full training set) ====

PART decision list
-----
referral_source = other AND
age_1 = '(-inf-91.8]' AND
FTI_2 = '(-inf-159.2]' AND
query_hypothyroid = f: negative (1878.08/16.0)

T3_1 = '(2.16-inf)': negative (656.23/7.04)

T3_measured = t AND
referral_source = SVHC AND
query_hypothyroid = f: negative (194.0/6.0)

T3_measured = t AND
thyroid_surgery = f AND
psych = f AND
TT4_1 = '(87.6-inf)' AND
TSH_measured = t AND
TT4_2 = '(-inf-173.2)': negative (585.16/97.0)

T3_measured = f: negative (126.02/1.96)

thyroid_surgery = f AND
T4U_measured = t AND
T4U_2 = '(1.078-inf)': negative (32.0/2.0)

thyroid_surgery = f AND
T4U_measured = f: negative (16.84/1.0)

thyroid_surgery = f AND
query_hyperthyroid = f AND
TSH_1 = '(-inf-106.004]' AND
on_thyroxine = f AND
sick = t AND
referral_source = SVI: sick (16.0/4.0)

thyroid_surgery = f AND
query_hyperthyroid = f AND
TSH_1 = '(-inf-106.004]' AND
on_thyroxine = f AND
referral_source = SVI AND
T4U_1 = '(0.664-inf)' AND
sex = F: negative (101.09/28.1)

thyroid_surgery = f AND
psych = f AND
query_hyperthyroid = f AND
TSH_1 = '(-inf-106.004]' AND
on_thyroxine = f AND
TT4_2 = '(-inf-173.2)' AND
referral_source = SVI AND
query_on_thyroxine = f AND
FTI_1 = '(80.6-inf)' AND
T4U_1 = '(0.664-inf)': negative (56.46/16.45)

thyroid_surgery = f AND
query_hyperthyroid = f AND
on_thyroxine = f AND
TSH_1 = '(-inf-106.004]' AND
referral_source = SVI AND
psych = f AND
sex = M: sick (35.05/12.72)

psych = f AND
thyroid_surgery = f AND
sex = M: negative (14.69/1.03)

sex = F AND

```

```

thyroid_surgery = f AND
query_hyperthyroid = f AND
FTI_1 = '(-inf-80.6]': negative (13.8/3.09)

sex = F AND
thyroid_surgery = t: negative (9.32)

query_hyperthyroid = f AND
sex = F AND
on_thyroxine = t: negative (6.0/1.0)

query_hyperthyroid = f AND
sex = F AND
referral_source = other AND
query_hypothyroid = f: sick (5.0)

sex = F AND
query_hyperthyroid = f AND
TT4_1 = '(-inf-87.6]' AND
psych = f: negative (14.22/6.0)

psych = f: negative (6.03/0.0)

: sick (6.0/1.0)

Number of Rules : 19

Time taken to build model: 0.15 seconds

==== Evaluation on test split ====

Time taken to test model on test split: 0 seconds

==== Summary ====

Correctly Classified Instances      1211          94.4618 %
Incorrectly Classified Instances    71           5.5382 %
Kappa statistic                   0.2189
Mean absolute error               0.086
Root mean squared error          0.2118
Relative absolute error          75.3706 %
Root relative squared error     90.2184 %
Total Number of Instances        1282

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.994     0.853     0.949     0.994     0.971     0.281     0.873     0.988   negative
      0.147     0.006     0.611     0.147     0.237     0.281     0.873     0.320     sick
Weighted Avg.   0.945     0.804     0.930     0.945     0.928     0.281     0.873     0.949

==== Confusion Matrix ====

      a      b  <-- classified as
1200    7 |    a = negative
      64   11 |    b = sick

```

## Leave one out fold

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B5-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  48
            age_1
            age_2
            age_3
            age_4
            sex
            on_thyroxine
            query_on_thyroxine
            on_antithyroid_medication
            sick
            pregnant
            thyroid_surgery
            i131_treatment
            query_hypothyroid
            query_hyperthyroid
            lithium
            goitre
            tumor
            hypopituitary

```

```

psych
TSH_measured
TSH_1
TSH_2
TSH_3
TSH_4
T3_measured
T3_1
T3_2
T3_3
T3_4
TT4_measured
TT4_1
TT4_2
TT4_3
TT4_4
T4U_measured
T4U_1
T4U_2
T4U_3
T4U_4
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
TBG_measured
TBG
referral_source
Class
Test mode: 3772-fold cross-validation

==== Classifier model (full training set) ====

PART decision list
-----
referral_source = other AND
age_1 = '(-inf-91.8]' AND
FTI_2 = '(-inf-159.2]' AND
query_hypothyroid = f: negative (1878.08/16.0)

T3_1 = '(2.16-inf)': negative (656.23/7.04)

T3_measured = t AND
referral_source = SVHC AND
query_hypothyroid = f: negative (194.0/6.0)

T3_measured = t AND
thyroid_surgery = f AND
psych = f AND
TT4_1 = '(87.6-inf)' AND
TSH_measured = t AND
TT4_2 = '(-inf-173.2)': negative (585.16/97.0)

T3_measured = f: negative (126.02/1.96)

thyroid_surgery = f AND
T4U_measured = t AND
T4U_2 = '(1.078-inf)': negative (32.0/2.0)

thyroid_surgery = f AND
T4U_measured = f: negative (16.84/1.0)

thyroid_surgery = f AND
query_hyperthyroid = f AND
TSH_1 = '(-inf-106.004]' AND
on_thyroxine = f AND
sick = t AND
referral_source = SVI: sick (16.0/4.0)

thyroid_surgery = f AND
query_hyperthyroid = f AND
TSH_1 = '(-inf-106.004]' AND
on_thyroxine = f AND
referral_source = SVI AND
T4U_1 = '(0.664-inf)' AND
sex = F: negative (101.09/28.1)

thyroid_surgery = f AND
psych = f AND
query_hyperthyroid = f AND
TSH_1 = '(-inf-106.004]' AND
on_thyroxine = f AND
TT4_2 = '(-inf-173.2)' AND
referral_source = SVI AND
query_on_thyroxine = f AND

```

```

FTI_1 = '(80.6-inf)' AND
T4U_1 = '(0.664-inf)': negative (56.46/16.45)

thyroid_surgery = f AND
query_hyperthyroid = f AND
on_thyroxine = f AND
TSH_1 = '(-inf-106.004]' AND
referral_source = SVI AND
psych = f AND
sex = M: sick (35.05/12.72)

psych = f AND
thyroid_surgery = f AND
sex = M: negative (14.69/1.03)

sex = F AND
thyroid_surgery = f AND
query_hyperthyroid = f AND
FTI_1 = '(-inf-80.6]': negative (13.8/3.09)

sex = F AND
thyroid_surgery = t: negative (9.32)

query_hyperthyroid = f AND
sex = F AND
on_thyroxine = t: negative (6.0/1.0)

query_hyperthyroid = f AND
sex = F AND
referral_source = other AND
query_hypothyroid = f: sick (5.0)

sex = F AND
query_hyperthyroid = f AND
TT4_1 = '(-inf-87.6]' AND
psych = f: negative (14.22/6.0)

psych = f: negative (6.03/0.0)

: sick (6.0/1.0)

Number of Rules : 19

Time taken to build model: 0.16 seconds

==== Stratified cross-validation ====
==== Summary ====

    Correctly Classified Instances      3549          94.088 %
    Incorrectly Classified Instances   223           5.912 %
    Kappa statistic                   0.2509
    Mean absolute error              0.0923
    Root mean squared error          0.221
    Relative absolute error          80.1343 %
    Root relative squared error     92.1383 %
    Total Number of Instances       3772

==== Detailed Accuracy By Class ====

          TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
        0.990     0.818     0.949     0.990     0.969     0.294     0.774     0.972   negative
        0.182     0.010     0.553     0.182     0.274     0.294     0.774     0.260     sick
    Weighted Avg.   0.941     0.769     0.925     0.941     0.927     0.294     0.774     0.929

==== Confusion Matrix ====

      a      b  <-- classified as
3507  34 |    a = negative
  189   42 |    b = sick

```

### 3.1.3 10 bins

#### Cross Validation 5 folds

```

==== Run information ====
Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B10-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  78
             age_1
             age_2

```

```

age_3
age_4
age_5
age_6
age_7
age_8
age_9
sex
on_thyroxine
query_on_thyroxine
on_antithyroid_medication
sick
pregnant
thyroid_surgery
I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH_1
TSH_2
TSH_3
TSH_4
TSH_5
TSH_6
TSH_7
TSH_8
TSH_9
T3_measured
T3_1
T3_2
T3_3
T3_4
T3_5
T3_6
T3_7
T3_8
T3_9
TT4_measured
TT4_1
TT4_2
TT4_3
TT4_4
TT4_5
TT4_6
TT4_7
TT4_8
TT4_9
T4U_measured
T4U_1
T4U_2
T4U_3
T4U_4
T4U_5
T4U_6
T4U_7
T4U_8
T4U_9
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
FTI_5
FTI_6
FTI_7
FTI_8
FTI_9
TBG_measured
TBG
referral_source
Class
Test mode: 5-fold cross-validation

==== Classifier model (full training set) ===

PART decision list
-----
T3_1 = '(1.105-inf)' AND
FTI_4 = '(-inf-159.2]': negative (3170.02/7.69)

referral_source = other AND

```

```

T3_measured = f: negative (118.6/0.1)

T3_1 = '(1.105-inf)' AND
T3_3 = '(3.215-inf)': negative (60.01/0.03)

T3_1 = '(1.105-inf)' AND
on_thyroxine = t: negative (34.86)

TT4_1 = '(-inf-44.8]' AND
TSH_1 = '(53.0045-inf)': negative (26.16)

T3_measured = t AND
T3_1 = '(1.105-inf)' AND
TT4_3 = '(-inf-130.4)': negative (13.44)

T3_measured = t AND
FTI_4 = '(-inf-159.2]' AND
TT4_1 = '(44.8-inf)' AND
referral_source = SVI AND
FTI_2 = '(80.6-inf)': sick (146.27/2.03)

sick = f AND
T3_measured = t AND
TT4_5 = '(-inf-216)' AND
T4U_1 = '(0.457-inf)' AND
FTI_1 = '(41.3-inf)' AND
T4U_4 = '(-inf-1.078]' AND
FTI_4 = '(159.2-inf)' AND
query_hypothyroid = f AND
TSH_measured = t AND
on_thyroxine = f AND
query_hyperthyroid = f AND
TT4_2 = '(87.6-inf)' AND
referral_source = SVI AND
T3_2 = '(-inf-2.16)': negative (23.79/4.79)

sick = f AND
T3_measured = t AND
TT4_5 = '(-inf-216)' AND
T4U_1 = '(0.457-inf)' AND
FTI_1 = '(-inf-41.3)': negative (12.58/1.13)

T3_measured = f: negative (10.18/0.18)

sick = t: sick (8.34)

TT4_5 = '(-inf-216)' AND
T4U_1 = '(0.457-inf)' AND
T4U_4 = '(-inf-1.078]' AND
TSH_measured = t AND
referral_source = SVHC AND
T3_1 = '(-inf-1.105)': sick (8.0)

psych = f AND
TT4_5 = '(-inf-216)' AND
T4U_1 = '(0.457-inf)' AND
T4U_4 = '(-inf-1.078]' AND
referral_source = other AND
T4U_measured = t AND
T3_2 = '(2.16-inf)': negative (9.0/2.0)

psych = f AND
TT4_5 = '(-inf-216)' AND
T4U_1 = '(-inf-0.457)': negative (5.28/0.06)

psych = f AND
TT4_5 = '(-inf-216)' AND
T4U_4 = '(1.078-inf)': negative (5.28/0.06)

psych = f AND
TT4_5 = '(-inf-216)' AND
referral_source = other AND
T4U_measured = t AND
TT4_4 = '(-inf-173.2]' AND
query_hypothyroid = t AND
FTI_2 = '(80.6-inf)': sick (4.0)

psych = f AND
TT4_5 = '(-inf-216)' AND
referral_source = other AND
T4U_measured = t AND
query_hypothyroid = f AND
TT4_4 = '(-inf-173.2]' AND
TSH_2 = '(-inf-106.004]' AND
T4U_2 = '(0.664-inf)' AND
T4U_3 = '(-inf-0.871]' AND
age_1 = '(-inf-46.4)': negative (5.96)

```

```

psych = f AND
TT4_5 = '(-inf-216]' AND
referral_source = other AND
T4U_measured = t AND
query_hypothyroid = f AND
TT4_4 = '(-inf-173.2]' AND
TSH_2 = '(-inf-106.004]' AND
T4U_2 = '(0.664-inf)' AND
sex = M: negative (11.47/3.26)

psych = f AND
TT4_5 = '(-inf-216]' AND
query_hypothyroid = t AND
T4U_3 = '(-inf-0.871]': negative (5.0)

psych = f AND
TT4_5 = '(-inf-216]' AND
referral_source = SVI AND
sex = F AND
age_1 = '(46.4-inf)' AND
T3_1 = '(-inf-1.105]': negative (18.67/6.83)

referral_source = SVI AND
TT4_4 = '(-inf-173.2]' AND
T3_1 = '(-inf-1.105]': sick (15.54/1.0)

T4U_measured = t AND
psych = f AND
TT4_5 = '(-inf-216]' AND
referral_source = SVI AND
age_1 = '(46.4-inf)': sick (5.0/1.0)

T3_2 = '(2.16-inf)': negative (7.65)

T4U_measured = t AND
psych = f AND
referral_source = other AND
query_hypothyroid = f AND
TT4_4 = '(-inf-173.2]' AND
TSH_2 = '(-inf-106.004]' AND
T4U_2 = '(0.664-inf)' AND
on_thyroxine = f AND
age_1 = '(-inf-46.4]' AND
T3_1 = '(-inf-1.105]': negative (4.73)

T4U_measured = t AND
psych = f AND
referral_source = other AND
query_hypothyroid = f AND
age_1 = '(46.4-inf)' AND
TT4_4 = '(-inf-173.2]' AND
TSH_1 = '(-inf-53.0045]' AND
FTI_4 = '(-inf-159.2]' AND
T4U_2 = '(0.664-inf)' AND
on_thyroxine = f: sick (10.09/3.05)

T4U_measured = t AND
psych = f AND
referral_source = other AND
query_hypothyroid = f AND
age_1 = '(46.4-inf)' AND
TT4_2 = '(87.6-inf)' AND
TT4_4 = '(-inf-173.2]': negative (8.74/1.69)

T4U_measured = t AND
psych = f AND
referral_source = other AND
query_hypothyroid = f: sick (11.01/0.01)

: negative (12.32/2.0)

Number of Rules : 28

Time taken to build model: 0.2 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3705           98.2238 %
Incorrectly Classified Instances    67            1.7762 %
Kappa statistic                   0.838
Mean absolute error                0.0235
Root mean squared error            0.1215
Relative absolute error             20.4155 %
Root relative squared error        50.6827 %

```

```
Total Number of Instances          3772
==== Detailed Accuracy By Class ====
      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.994     0.195     0.987     0.994     0.991     0.839     0.964     0.994     negative
      0.805     0.006     0.894     0.805     0.847     0.839     0.964     0.886     sick
Weighted Avg.    0.982     0.183     0.982     0.982     0.982     0.839     0.964     0.987
==== Confusion Matrix ====
      a     b  <-- classified as
3519  22 |     a = negative
     45 186 |     b = sick
```

## Cross Validation 10 folds

```
==== Run information ====
Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B10-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  78
              age_1
              age_2
              age_3
              age_4
              age_5
              age_6
              age_7
              age_8
              age_9
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              TSH_2
              TSH_3
              TSH_4
              TSH_5
              TSH_6
              TSH_7
              TSH_8
              TSH_9
              T3_measured
              T3_1
              T3_2
              T3_3
              T3_4
              T3_5
              T3_6
              T3_7
              T3_8
              T3_9
              TT4_measured
              TT4_1
              TT4_2
              TT4_3
              TT4_4
              TT4_5
              TT4_6
              TT4_7
              TT4_8
              TT4_9
              T4U_measured
              T4U_1
              T4U_2
              T4U_3
              T4U_4
              T4U_5
              T4U_6
```

```

T4U_7
T4U_8
T4U_9
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
FTI_5
FTI_6
FTI_7
FTI_8
FTI_9
TBG_measured
TBG
referral_source
Class
Test mode: 10-fold cross-validation

==== Classifier model (full training set) ====

PART decision list
-----
T3_1 = '(1.105-inf)' AND
FTI_4 = '(-inf-159.2)': negative (3170.02/7.69)

referral_source = other AND
T3_measured = f: negative (118.6/0.1)

T3_1 = '(1.105-inf)' AND
T3_3 = '(3.215-inf)': negative (60.01/0.03)

T3_1 = '(1.105-inf)' AND
on_thyroxine = t: negative (34.86)

TT4_1 = '(-inf-44.8)' AND
TSH_1 = '(53.0045-inf)': negative (26.16)

T3_measured = t AND
T3_1 = '(1.105-inf)' AND
TT4_3 = '(-inf-130.4)': negative (13.44)

T3_measured = t AND
FTI_4 = '(-inf-159.2)' AND
TT4_1 = '(44.8-inf)' AND
referral_source = SVI AND
FTI_2 = '(80.6-inf)': sick (146.27/2.03)

sick = f AND
T3_measured = t AND
TT4_5 = '(-inf-216)' AND
T4U_1 = '(0.457-inf)' AND
FTI_1 = '(41.3-inf)' AND
T4U_4 = '(-inf-1.078)' AND
FTI_4 = '(159.2-inf)' AND
query_hypothyroid = f AND
TSH_measured = t AND
on_thyroxine = f AND
query_hyperthyroid = f AND
TT4_2 = '(87.6-inf)' AND
referral_source = SVI AND
T3_2 = '(-inf-2.16)': negative (23.79/4.79)

sick = f AND
T3_measured = t AND
TT4_5 = '(-inf-216)' AND
T4U_1 = '(0.457-inf)' AND
FTI_1 = '(-inf-41.3)': negative (12.58/1.13)

T3_measured = f: negative (10.18/0.18)

sick = t: sick (8.34)

TT4_5 = '(-inf-216)' AND
T4U_1 = '(0.457-inf)' AND
T4U_4 = '(-inf-1.078)' AND
TSH_measured = t AND
referral_source = SVHC AND
T3_1 = '(-inf-1.105)': sick (8.0)

psych = f AND
TT4_5 = '(-inf-216)' AND
T4U_1 = '(0.457-inf)' AND
T4U_4 = '(-inf-1.078)' AND
referral_source = other AND
T4U_measured = t AND

```

```

T3_2 = '(2.16-inf)': negative (9.0/2.0)

psych = f AND
TT4_5 = '(-inf-216]' AND
T4U_1 = '(-inf-0.457]': negative (5.28/0.06)

psych = f AND
TT4_5 = '(-inf-216]' AND
T4U_4 = '(1.078-inf)': negative (5.28/0.06)

psych = f AND
TT4_5 = '(-inf-216]' AND
referral_source = other AND
T4U_measured = t AND
TT4_4 = '(-inf-173.2]' AND
query_hypothyroid = t AND
FTI_2 = '(80.6-inf)': sick (4.0)

psych = f AND
TT4_5 = '(-inf-216]' AND
referral_source = other AND
T4U_measured = t AND
query_hypothyroid = f AND
TT4_4 = '(-inf-173.2]' AND
TSH_2 = '(-inf-106.004]' AND
T4U_2 = '(0.664-inf)' AND
T4U_3 = '(-inf-0.871]' AND
age_1 = '(-inf-46.4]': negative (5.96)

psych = f AND
TT4_5 = '(-inf-216]' AND
referral_source = other AND
T4U_measured = t AND
query_hypothyroid = f AND
TT4_4 = '(-inf-173.2]' AND
TSH_2 = '(-inf-106.004]' AND
T4U_2 = '(0.664-inf)' AND
sex = M: negative (11.47/3.26)

psych = f AND
TT4_5 = '(-inf-216]' AND
query_hypothyroid = t AND
T4U_3 = '(-inf-0.871]': negative (5.0)

psych = f AND
TT4_5 = '(-inf-216]' AND
referral_source = SVI AND
sex = F AND
age_1 = '(46.4-inf)' AND
T3_1 = '(-inf-1.105]': negative (18.67/6.83)

referral_source = SVI AND
TT4_4 = '(-inf-173.2]' AND
T3_1 = '(-inf-1.105]': sick (15.54/1.0)

T4U_measured = t AND
psych = f AND
TT4_5 = '(-inf-216]' AND
referral_source = SVI AND
age_1 = '(46.4-inf)': sick (5.0/1.0)

T3_2 = '(2.16-inf)': negative (7.65)

T4U_measured = t AND
psych = f AND
referral_source = other AND
query_hypothyroid = f AND
TT4_4 = '(-inf-173.2]' AND
TSH_2 = '(-inf-106.004]' AND
T4U_2 = '(0.664-inf)' AND
on_thyroxine = f AND
age_1 = '(-inf-46.4]': negative (4.73)

T4U_measured = t AND
psych = f AND
referral_source = other AND
query_hypothyroid = f AND
age_1 = '(46.4-inf)' AND
TT4_4 = '(-inf-173.2]' AND
TSH_1 = '(-inf-53.0045]' AND
FTI_4 = '(-inf-159.2]' AND
T4U_2 = '(0.664-inf)' AND
on_thyroxine = f: sick (10.09/3.05)

T4U_measured = t AND
psych = f AND

```

```

referral_source = other AND
query_hypothyroid = f AND
age_1 = '(46.4-inf)' AND
TT4_2 = '(87.6-inf)' AND
TT4_4 = '(-inf-173.2]': negative (8.74/1.69)

T4U_measured = t AND
psych = f AND
referral_source = other AND
query_hypothyroid = f: sick (11.01/0.01)

: negative (12.32/2.0)

Number of Rules : 28

Time taken to build model: 0.23 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3694           97.9321 %
Incorrectly Classified Instances   78            2.0679 %
Kappa statistic                   0.8157
Mean absolute error               0.026
Root mean squared error          0.1289
Relative absolute error          22.5511 %
Root relative squared error     53.7485 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.991     0.195     0.987     0.991     0.989     0.816     0.947     0.989     negative
      0.805     0.009     0.849     0.805     0.827     0.816     0.947     0.847     sick
Weighted Avg.       0.979     0.183     0.979     0.979     0.979     0.816     0.947     0.980

==== Confusion Matrix ====

      a      b  <-- classified as
3508    33 |    a = negative
      45   186 |    b = sick

```

## Cross Validation 15 folds

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B10-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  78
              age_1
              age_2
              age_3
              age_4
              age_5
              age_6
              age_7
              age_8
              age_9
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              i131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              TSH_2
              TSH_3
              TSH_4
              TSH_5
              TSH_6
              TSH_7
              TSH_8

```

```

TSH_9
T3_measured
T3_1
T3_2
T3_3
T3_4
T3_5
T3_6
T3_7
T3_8
T3_9
TT4_measured
TT4_1
TT4_2
TT4_3
TT4_4
TT4_5
TT4_6
TT4_7
TT4_8
TT4_9
T4U_measured
T4U_1
T4U_2
T4U_3
T4U_4
T4U_5
T4U_6
T4U_7
T4U_8
T4U_9
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
FTI_5
FTI_6
FTI_7
FTI_8
FTI_9
TBG_measured
TBG
referral_source
Class
Test mode: 15-fold cross-validation

==== Classifier model (full training set) ====
PART decision list
-----
T3_1 = '(1.105-inf)' AND
FTI_4 = '(-inf-159.2]': negative (3170.02/7.69)

referral_source = other AND
T3_measured = f: negative (118.6/0.1)

T3_1 = '(1.105-inf)' AND
T3_3 = '(3.215-inf)': negative (60.01/0.03)

T3_1 = '(1.105-inf)' AND
on_thyroxine = t: negative (34.86)

TT4_1 = '(-inf-44.8]' AND
TSH_1 = '(53.0045-inf)': negative (26.16)

T3_measured = t AND
T3_1 = '(1.105-inf)' AND
TT4_3 = '(-inf-130.4]': negative (13.44)

T3_measured = t AND
FTI_4 = '(-inf-159.2]' AND
TT4_1 = '(44.8-inf)' AND
referral_source = SVI AND
FTI_2 = '(80.6-inf)': sick (146.27/2.03)

sick = f AND
T3_measured = t AND
TT4_5 = '(-inf-216]' AND
T4U_1 = '(0.457-inf)' AND
FTI_1 = '(41.3-inf)' AND
T4U_4 = '(-inf-1.078]' AND
FTI_4 = '(159.2-inf)' AND
query_hypothyroid = f AND
TSH_measured = t AND
on_thyroxine = f AND

```

```

query_hyperthyroid = f AND
TT4_2 = '(87.6-inf)' AND
referral_source = SVI AND
T3_2 = '(-inf-2.16]': negative (23.79/4.79)

sick = f AND
T3_measured = t AND
TT4_5 = '(-inf-216]' AND
T4U_1 = '(0.457-inf)' AND
FTI_1 = '(-inf-41.3]': negative (12.58/1.13)

T3_measured = f: negative (10.18/0.18)

sick = t: sick (8.34)

TT4_5 = '(-inf-216]' AND
T4U_1 = '(0.457-inf)' AND
T4U_4 = '(-inf-1.078]' AND
TSH_measured = t AND
referral_source = SVHC AND
T3_1 = '(-inf-1.105]': sick (8.0)

psych = f AND
TT4_5 = '(-inf-216]' AND
T4U_1 = '(0.457-inf)' AND
T4U_4 = '(-inf-1.078]' AND
referral_source = other AND
T4U_measured = t AND
T3_2 = '(2.16-inf)': negative (9.0/2.0)

psych = f AND
TT4_5 = '(-inf-216]' AND
T4U_1 = '(-inf-0.457]': negative (5.28/0.06)

psych = f AND
TT4_5 = '(-inf-216]' AND
T4U_4 = '(1.078-inf)': negative (5.28/0.06)

psych = f AND
TT4_5 = '(-inf-216]' AND
referral_source = other AND
T4U_measured = t AND
TT4_4 = '(-inf-173.2]' AND
query_hypothyroid = t AND
FTI_2 = '(80.6-inf)': sick (4.0)

psych = f AND
TT4_5 = '(-inf-216]' AND
referral_source = other AND
T4U_measured = t AND
query_hypothyroid = f AND
TT4_4 = '(-inf-173.2]' AND
TSH_2 = '(-inf-106.004]' AND
T4U_2 = '(0.664-inf)' AND
T4U_3 = '(-inf-0.871]' AND
age_1 = '(-inf-46.4]': negative (5.96)

psych = f AND
TT4_5 = '(-inf-216]' AND
referral_source = other AND
T4U_measured = t AND
query_hypothyroid = f AND
TT4_4 = '(-inf-173.2]' AND
TSH_2 = '(-inf-106.004]' AND
T4U_2 = '(0.664-inf)' AND
sex = M: negative (11.47/3.26)

psych = f AND
TT4_5 = '(-inf-216]' AND
query_hypothyroid = t AND
T4U_3 = '(-inf-0.871]': negative (5.0)

psych = f AND
TT4_5 = '(-inf-216]' AND
referral_source = SVI AND
sex = F AND
age_1 = '(46.4-inf)' AND
T3_1 = '(-inf-1.105]': negative (18.67/6.83)

referral_source = SVI AND
TT4_4 = '(-inf-173.2]' AND
T3_1 = '(-inf-1.105]': sick (15.54/1.0)

T4U_measured = t AND
psych = f AND
TT4_5 = '(-inf-216]' AND
referral_source = SVI AND

```

```

age_1 = '(46.4-inf)': sick (5.0/1.0)
T3_2 = '(2.16-inf)': negative (7.65)

T4U_measured = t AND
psych = f AND
referral_source = other AND
query_hypothyroid = f AND
TT4_4 = '(-inf-173.2]' AND
TSH_2 = '(-inf-106.004]' AND
T4U_2 = '(0.664-inf)' AND
on_thyroxine = f AND
age_1 = '(-inf-46.4]' AND
T3_1 = '(-inf-1.105]': negative (4.73)

T4U_measured = t AND
psych = f AND
referral_source = other AND
query_hypothyroid = f AND
age_1 = '(46.4-inf)' AND
TT4_4 = '(-inf-173.2]' AND
TSH_1 = '(-inf-53.0045]' AND
FTI_4 = '(-inf-159.2]' AND
T4U_2 = '(0.664-inf)' AND
on_thyroxine = f: sick (10.09/3.05)

T4U_measured = t AND
psych = f AND
referral_source = other AND
query_hypothyroid = f AND
age_1 = '(46.4-inf)' AND
TT4_2 = '(87.6-inf)' AND
TT4_4 = '(-inf-173.2]': negative (8.74/1.69)

T4U_measured = t AND
psych = f AND
referral_source = other AND
query_hypothyroid = f: sick (11.01/0.01)

: negative (12.32/2.0)

Number of Rules : 28

Time taken to build model: 0.2 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3701      98.1177 %
Incorrectly Classified Instances    71       1.8823 %
Kappa statistic                      0.8312
Mean absolute error                  0.0248
Root mean squared error              0.1252
Relative absolute error              21.4965 %
Root relative squared error         52.1988 %
Total Number of Instances           3772

==== Detailed Accuracy By Class ====

          TP Rate   FP Rate   Precision   Recall   F-Measure   MCC     ROC Area   PRC Area   Class
0.992      0.186      0.988      0.992      0.990      0.832      0.945      0.991   negative
0.814      0.008      0.870      0.814      0.841      0.832      0.945      0.837   sick
Weighted Avg.   0.981      0.175      0.981      0.981      0.981      0.832      0.945      0.981

==== Confusion Matrix ====

      a      b  <-- classified as
3513    28 |     a = negative
      43   188 |     b = sick

```

## Percentage Split 50%

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B10-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  78
              age_1
              age_2
              age_3
              age_4
              age_5

```

```

age_6
age_7
age_8
age_9
sex
on_thyroxine
query_on_thyroxine
on_antithyroid_medication
sick
pregnant
thyroid_surgery
I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH_1
TSH_2
TSH_3
TSH_4
TSH_5
TSH_6
TSH_7
TSH_8
TSH_9
T3_measured
T3_1
T3_2
T3_3
T3_4
T3_5
T3_6
T3_7
T3_8
T3_9
TT4_measured
TT4_1
TT4_2
TT4_3
TT4_4
TT4_5
TT4_6
TT4_7
TT4_8
TT4_9
T4U_measured
T4U_1
T4U_2
T4U_3
T4U_4
T4U_5
T4U_6
T4U_7
T4U_8
T4U_9
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
FTI_5
FTI_6
FTI_7
FTI_8
FTI_9
TBG_measured
TBG
referral_source
Class
Test mode: split 50.0% train, remainder test

==== Classifier model (full training set) ====

PART decision list
-----
T3_1 = '(1.105-inf)' AND
FTI_4 = '(-inf-159.2)': negative (3170.02/7.69)

referral_source = other AND
T3_measured = f: negative (118.6/0.1)

T3_1 = '(1.105-inf)' AND

```

```

T3_3 = '(3.215-inf)': negative (60.01/0.03)
T3_1 = '(1.105-inf)' AND
on_thyroxine = t: negative (34.86)

TT4_1 = '(-inf-44.8]' AND
TSH_1 = '(53.0045-inf)': negative (26.16)

T3_measured = t AND
T3_1 = '(1.105-inf)' AND
TT4_3 = '(-inf-130.4)': negative (13.44)

T3_measured = t AND
FTI_4 = '(-inf-159.2]' AND
TT4_1 = '(44.8-inf)' AND
referral_source = SVI AND
FTI_2 = '(80.6-inf)': sick (146.27/2.03)

sick = f AND
T3_measured = t AND
TT4_5 = '(-inf-216]' AND
T4U_1 = '(0.457-inf)' AND
FTI_1 = '(41.3-inf)' AND
T4U_4 = '(-inf-1.078]' AND
FTI_4 = '(159.2-inf)' AND
query_hypothyroid = f AND
TSH_measured = t AND
on_thyroxine = f AND
query_hyperthyroid = f AND
TT4_2 = '(87.6-inf)' AND
referral_source = SVI AND
T3_2 = '(-inf-2.16)': negative (23.79/4.79)

sick = f AND
T3_measured = t AND
TT4_5 = '(-inf-216]' AND
T4U_1 = '(0.457-inf)' AND
FTI_1 = '(-inf-41.3)': negative (12.58/1.13)

T3_measured = f: negative (10.18/0.18)

sick = t: sick (8.34)

TT4_5 = '(-inf-216]' AND
T4U_1 = '(0.457-inf)' AND
T4U_4 = '(-inf-1.078]' AND
TSH_measured = t AND
referral_source = SVHC AND
T3_1 = '(-inf-1.105)': sick (8.0)

psych = f AND
TT4_5 = '(-inf-216]' AND
T4U_1 = '(0.457-inf)' AND
T4U_4 = '(-inf-1.078]' AND
referral_source = other AND
T4U_measured = t AND
T3_2 = '(2.16-inf)': negative (9.0/2.0)

psych = f AND
TT4_5 = '(-inf-216]' AND
T4U_1 = '(-inf-0.457)': negative (5.28/0.06)

psych = f AND
TT4_5 = '(-inf-216]' AND
T4U_4 = '(1.078-inf)': negative (5.28/0.06)

psych = f AND
TT4_5 = '(-inf-216]' AND
referral_source = other AND
T4U_measured = t AND
TT4_4 = '(-inf-173.2]' AND
query_hypothyroid = t AND
FTI_2 = '(80.6-inf)': sick (4.0)

psych = f AND
TT4_5 = '(-inf-216]' AND
referral_source = other AND
T4U_measured = t AND
query_hypothyroid = f AND
TT4_4 = '(-inf-173.2]' AND
TSH_2 = '(-inf-106.004]' AND
T4U_2 = '(0.664-inf)' AND
T4U_3 = '(-inf-0.871]' AND
age_1 = '(-inf-46.4)': negative (5.96)

psych = f AND
TT4_5 = '(-inf-216]' AND

```

```

referral_source = other AND
T4U_measured = t AND
query_hypothyroid = f AND
TT4_4 = '(-inf-173.2]' AND
TSH_2 = '(-inf-106.004]' AND
T4U_2 = '(0.664-inf)' AND
sex = M: negative (11.47/3.26)

psych = f AND
TT4_5 = '(-inf-216]' AND
query_hypothyroid = t AND
T4U_3 = '(-inf-0.871]': negative (5.0)

psych = f AND
TT4_5 = '(-inf-216]' AND
referral_source = SVI AND
sex = F AND
age_1 = '(46.4-inf)' AND
T3_1 = '(-inf-1.105]': negative (18.67/6.83)

referral_source = SVI AND
TT4_4 = '(-inf-173.2]' AND
T3_1 = '(-inf-1.105]': sick (15.54/1.0)

T4U_measured = t AND
psych = f AND
TT4_5 = '(-inf-216]' AND
referral_source = SVI AND
age_1 = '(46.4-inf)': sick (5.0/1.0)

T3_2 = '(2.16-inf)': negative (7.65)

T4U_measured = t AND
psych = f AND
referral_source = other AND
query_hypothyroid = f AND
TT4_4 = '(-inf-173.2]' AND
TSH_2 = '(-inf-106.004]' AND
T4U_2 = '(0.664-inf)' AND
on_thyroxine = f AND
age_1 = '(-inf-46.4]' AND
T3_1 = '(-inf-1.105]': negative (4.73)

T4U_measured = t AND
psych = f AND
referral_source = other AND
query_hypothyroid = f AND
age_1 = '(46.4-inf)' AND
TT4_4 = '(-inf-173.2]' AND
TSH_1 = '(-inf-53.0045]' AND
FTI_4 = '(-inf-159.2]' AND
T4U_2 = '(0.664-inf)' AND
on_thyroxine = f: sick (10.09/3.05)

T4U_measured = t AND
psych = f AND
referral_source = other AND
query_hypothyroid = f AND
age_1 = '(46.4-inf)' AND
TT4_2 = '(87.6-inf)' AND
TT4_4 = '(-inf-173.2]': negative (8.74/1.69)

T4U_measured = t AND
psych = f AND
referral_source = other AND
query_hypothyroid = f: sick (11.01/0.01)

: negative (12.32/2.0)

Number of Rules : 28

Time taken to build model: 0.23 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0.04 seconds

==== Summary ===

Correctly Classified Instances      1838          97.4549 %
Incorrectly Classified Instances    48           2.5451 %
Kappa statistic                   0.7864
Mean absolute error                0.0278
Root mean squared error            0.1391
Relative absolute error             24.061   %
Root relative squared error        57.6837 %

```

```
Total Number of Instances          1886
== Detailed Accuracy By Class ==
      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.985     0.179     0.988     0.985     0.986     0.787     0.941     0.990   negative
      0.821     0.015     0.780     0.821     0.800     0.787     0.941     0.799   sick
Weighted Avg.    0.975     0.169     0.975     0.975     0.975     0.787     0.941     0.978

== Confusion Matrix ==
      a     b  <-- classified as
1742  27 |     a = negative
   21  96 |     b = sick
```

## Percentage Split 60%

```
== Run information ==
Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B10-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  78
              age_1
              age_2
              age_3
              age_4
              age_5
              age_6
              age_7
              age_8
              age_9
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              TSH_2
              TSH_3
              TSH_4
              TSH_5
              TSH_6
              TSH_7
              TSH_8
              TSH_9
              T3_measured
              T3_1
              T3_2
              T3_3
              T3_4
              T3_5
              T3_6
              T3_7
              T3_8
              T3_9
              TT4_measured
              TT4_1
              TT4_2
              TT4_3
              TT4_4
              TT4_5
              TT4_6
              TT4_7
              TT4_8
              TT4_9
              T4U_measured
              T4U_1
              T4U_2
              T4U_3
              T4U_4
              T4U_5
              T4U_6
```

```

T4U_7
T4U_8
T4U_9
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
FTI_5
FTI_6
FTI_7
FTI_8
FTI_9
TBG_measured
TBG
referral_source
Class
Test mode: split 60.0% train, remainder test

==== Classifier model (full training set) ====

PART decision list
-----
T3_1 = '(1.105-inf)' AND
FTI_4 = '(-inf-159.2)': negative (3170.02/7.69)

referral_source = other AND
T3_measured = f: negative (118.6/0.1)

T3_1 = '(1.105-inf)' AND
T3_3 = '(3.215-inf)': negative (60.01/0.03)

T3_1 = '(1.105-inf)' AND
on_thyroxine = t: negative (34.86)

TT4_1 = '(-inf-44.8)' AND
TSH_1 = '(53.0045-inf)': negative (26.16)

T3_measured = t AND
T3_1 = '(1.105-inf)' AND
TT4_3 = '(-inf-130.4)': negative (13.44)

T3_measured = t AND
FTI_4 = '(-inf-159.2)' AND
TT4_1 = '(44.8-inf)' AND
referral_source = SVI AND
FTI_2 = '(80.6-inf)': sick (146.27/2.03)

sick = f AND
T3_measured = t AND
TT4_5 = '(-inf-216)' AND
T4U_1 = '(0.457-inf)' AND
FTI_1 = '(41.3-inf)' AND
T4U_4 = '(-inf-1.078)' AND
FTI_4 = '(159.2-inf)' AND
query_hypothyroid = f AND
TSH_measured = t AND
on_thyroxine = f AND
query_hyperthyroid = f AND
TT4_2 = '(87.6-inf)' AND
referral_source = SVI AND
T3_2 = '(-inf-2.16)': negative (23.79/4.79)

sick = f AND
T3_measured = t AND
TT4_5 = '(-inf-216)' AND
T4U_1 = '(0.457-inf)' AND
FTI_1 = '(-inf-41.3)': negative (12.58/1.13)

T3_measured = f: negative (10.18/0.18)

sick = t: sick (8.34)

TT4_5 = '(-inf-216)' AND
T4U_1 = '(0.457-inf)' AND
T4U_4 = '(-inf-1.078)' AND
TSH_measured = t AND
referral_source = SVHC AND
T3_1 = '(-inf-1.105)': sick (8.0)

psych = f AND
TT4_5 = '(-inf-216)' AND
T4U_1 = '(0.457-inf)' AND
T4U_4 = '(-inf-1.078)' AND
referral_source = other AND
T4U_measured = t AND

```

```

T3_2 = '(2.16-inf)': negative (9.0/2.0)

psych = f AND
TT4_5 = '(-inf-216]' AND
T4U_1 = '(-inf-0.457]': negative (5.28/0.06)

psych = f AND
TT4_5 = '(-inf-216]' AND
T4U_4 = '(1.078-inf)': negative (5.28/0.06)

psych = f AND
TT4_5 = '(-inf-216]' AND
referral_source = other AND
T4U_measured = t AND
TT4_4 = '(-inf-173.2]' AND
query_hypothyroid = t AND
FTI_2 = '(80.6-inf)': sick (4.0)

psych = f AND
TT4_5 = '(-inf-216]' AND
referral_source = other AND
T4U_measured = t AND
query_hypothyroid = f AND
TT4_4 = '(-inf-173.2]' AND
TSH_2 = '(-inf-106.004]' AND
T4U_2 = '(0.664-inf)' AND
T4U_3 = '(-inf-0.871]' AND
age_1 = '(-inf-46.4]': negative (5.96)

psych = f AND
TT4_5 = '(-inf-216]' AND
referral_source = other AND
T4U_measured = t AND
query_hypothyroid = f AND
TT4_4 = '(-inf-173.2]' AND
TSH_2 = '(-inf-106.004]' AND
T4U_2 = '(0.664-inf)' AND
sex = M: negative (11.47/3.26)

psych = f AND
TT4_5 = '(-inf-216]' AND
query_hypothyroid = t AND
T4U_3 = '(-inf-0.871]': negative (5.0)

psych = f AND
TT4_5 = '(-inf-216]' AND
referral_source = SVI AND
sex = F AND
age_1 = '(46.4-inf)' AND
T3_1 = '(-inf-1.105]': negative (18.67/6.83)

referral_source = SVI AND
TT4_4 = '(-inf-173.2]' AND
T3_1 = '(-inf-1.105]': sick (15.54/1.0)

T4U_measured = t AND
psych = f AND
TT4_5 = '(-inf-216]' AND
referral_source = SVI AND
age_1 = '(46.4-inf)': sick (5.0/1.0)

T3_2 = '(2.16-inf)': negative (7.65)

T4U_measured = t AND
psych = f AND
referral_source = other AND
query_hypothyroid = f AND
TT4_4 = '(-inf-173.2]' AND
TSH_2 = '(-inf-106.004]' AND
T4U_2 = '(0.664-inf)' AND
on_thyroxine = f AND
age_1 = '(-inf-46.4]': negative (4.73)

T4U_measured = t AND
psych = f AND
referral_source = other AND
query_hypothyroid = f AND
age_1 = '(46.4-inf)' AND
TT4_4 = '(-inf-173.2]' AND
TSH_1 = '(-inf-53.0045]' AND
FTI_4 = '(-inf-159.2]' AND
T4U_2 = '(0.664-inf)' AND
on_thyroxine = f: sick (10.09/3.05)

T4U_measured = t AND
psych = f AND

```

```

referral_source = other AND
query_hypothyroid = f AND
age_1 = '(46.4-inf)' AND
TT4_2 = '(87.6-inf)' AND
TT4_4 = '(-inf-173.2]': negative (8.74/1.69)

T4U_measured = t AND
psych = f AND
referral_source = other AND
query_hypothyroid = f: sick (11.01/0.01)

: negative (12.32/2.0)

Number of Rules : 28

Time taken to build model: 0.32 seconds

==== Evaluation on test split ====

Time taken to test model on test split: 0.02 seconds

==== Summary ====

Correctly Classified Instances      1480          98.0782 %
Incorrectly Classified Instances   29           1.9218 %
Kappa statistic                   0.812
Mean absolute error               0.0267
Root mean squared error          0.122
Relative absolute error          23.3251 %
Root relative squared error     51.7595 %
Total Number of Instances        1509

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC    ROC Area   PRC Area   Class
      0.995     0.247     0.985     0.995     0.990     0.816     0.958     0.993   negative
      0.753     0.005     0.905     0.753     0.822     0.816     0.958     0.875   sick
Weighted Avg.      0.981     0.233     0.980     0.981     0.980     0.816     0.958     0.986

==== Confusion Matrix ====

      a      b  <-- classified as
1413    7 |    a = negative
      22   67 |    b = sick

```

## Percentage Split 66%

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B10-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  78
              age_1
              age_2
              age_3
              age_4
              age_5
              age_6
              age_7
              age_8
              age_9
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              TSH_2
              TSH_3
              TSH_4
              TSH_5

```

```

TSH_6
TSH_7
TSH_8
TSH_9
T3_measured
T3_1
T3_2
T3_3
T3_4
T3_5
T3_6
T3_7
T3_8
T3_9
TT4_measured
TT4_1
TT4_2
TT4_3
TT4_4
TT4_5
TT4_6
TT4_7
TT4_8
TT4_9
T4U_measured
T4U_1
T4U_2
T4U_3
T4U_4
T4U_5
T4U_6
T4U_7
T4U_8
T4U_9
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
FTI_5
FTI_6
FTI_7
FTI_8
FTI_9
TBG_measured
TBG
referral_source
Class
Test mode: split 66.0% train, remainder test

==== Classifier model (full training set) ====

PART decision list
-----
T3_1 = '(1.105-inf)' AND
FTI_4 = '(-inf-159.2]': negative (3170.02/7.69)

referral_source = other AND
T3_measured = f: negative (118.6/0.1)

T3_1 = '(1.105-inf)' AND
T3_3 = '(3.215-inf)': negative (60.01/0.03)

T3_1 = '(1.105-inf)' AND
on_thyroxine = t: negative (34.86)

TT4_1 = '(-inf-44.8]' AND
TSH_1 = '(53.0045-inf)': negative (26.16)

T3_measured = t AND
T3_1 = '(1.105-inf)' AND
TT4_3 = '(-inf-130.4]': negative (13.44)

T3_measured = t AND
FTI_4 = '(-inf-159.2)' AND
TT4_1 = '(44.8-inf)' AND
referral_source = SVI AND
FTI_2 = '(80.6-inf)': sick (146.27/2.03)

sick = f AND
T3_measured = t AND
TT4_5 = '(-inf-216)' AND
T4U_1 = '(0.457-inf)' AND
FTI_1 = '(41.3-inf)' AND
T4U_4 = '(-inf-1.078)' AND
FTI_4 = '(159.2-inf)' AND

```

```

query_hypothyroid = f AND
TSH_measured = t AND
on_thyroxine = f AND
query_hyperthyroid = f AND
TT4_2 = '(87.6-inf)' AND
referral_source = SVI AND
T3_2 = '(-inf-2.16]': negative (23.79/4.79)

sick = f AND
T3_measured = t AND
TT4_5 = '(-inf-216]' AND
T4U_1 = '(0.457-inf)' AND
FTI_1 = '(-inf-41.3]': negative (12.58/1.13)

T3_measured = f: negative (10.18/0.18)

sick = t: sick (8.34)

TT4_5 = '(-inf-216]' AND
T4U_1 = '(0.457-inf)' AND
T4U_4 = '(-inf-1.078]' AND
TSH_measured = t AND
referral_source = SVHC AND
T3_1 = '(-inf-1.105]': sick (8.0)

psych = f AND
TT4_5 = '(-inf-216]' AND
T4U_1 = '(0.457-inf)' AND
T4U_4 = '(-inf-1.078]' AND
referral_source = other AND
T4U_measured = t AND
T3_2 = '(2.16-inf)': negative (9.0/2.0)

psych = f AND
TT4_5 = '(-inf-216]' AND
T4U_1 = '(-inf-0.457]': negative (5.28/0.06)

psych = f AND
TT4_5 = '(-inf-216]' AND
T4U_4 = '(1.078-inf)': negative (5.28/0.06)

psych = f AND
TT4_5 = '(-inf-216]' AND
referral_source = other AND
T4U_measured = t AND
TT4_4 = '(-inf-173.2]' AND
query_hypothyroid = t AND
FTI_2 = '(80.6-inf)': sick (4.0)

psych = f AND
TT4_5 = '(-inf-216]' AND
referral_source = other AND
T4U_measured = t AND
query_hypothyroid = f AND
TT4_4 = '(-inf-173.2]' AND
TSH_2 = '(-inf-106.004]' AND
T4U_2 = '(0.664-inf)' AND
T4U_3 = '(-inf-0.871]' AND
age_1 = '(-inf-46.4]': negative (5.96)

psych = f AND
TT4_5 = '(-inf-216]' AND
referral_source = other AND
T4U_measured = t AND
query_hypothyroid = f AND
TT4_4 = '(-inf-173.2]' AND
TSH_2 = '(-inf-106.004]' AND
T4U_2 = '(0.664-inf)' AND
sex = M: negative (11.47/3.26)

psych = f AND
TT4_5 = '(-inf-216]' AND
query_hypothyroid = t AND
T4U_3 = '(-inf-0.871]': negative (5.0)

psych = f AND
TT4_5 = '(-inf-216]' AND
referral_source = SVI AND
sex = F AND
age_1 = '(46.4-inf)' AND
T3_1 = '(-inf-1.105]': negative (18.67/6.83)

referral_source = SVI AND
TT4_4 = '(-inf-173.2]' AND
T3_1 = '(-inf-1.105]': sick (15.54/1.0)

T4U_measured = t AND

```

```

psych = f AND
TT4_5 = '(-inf-216]' AND
referral_source = SVI AND
age_1 = '(46.4-inf)': sick (5.0/1.0)

T3_2 = '(2.16-inf)': negative (7.65)

T4U_measured = t AND
psych = f AND
referral_source = other AND
query_hypothyroid = f AND
TT4_4 = '(-inf-173.2]' AND
TSH_2 = '(0.664-inf)' AND
T4U_2 = 'f' AND
on_thyroxine = f AND
age_1 = '(-inf-46.4]' AND
T3_1 = '(-inf-1.105]': negative (4.73)

T4U_measured = t AND
psych = f AND
referral_source = other AND
query_hypothyroid = f AND
age_1 = '(46.4-inf)' AND
TT4_4 = '(-inf-173.2]' AND
TSH_1 = '(-inf-53.0045]' AND
FTI_4 = '(-inf-159.2]' AND
T4U_2 = '(0.664-inf)' AND
on_thyroxine = f: sick (10.09/3.05)

T4U_measured = t AND
psych = f AND
referral_source = other AND
query_hypothyroid = f AND
age_1 = '(46.4-inf)' AND
TT4_2 = '(87.6-inf)' AND
TT4_4 = '(-inf-173.2]': negative (8.74/1.69)

T4U_measured = t AND
psych = f AND
referral_source = other AND
query_hypothyroid = f: sick (11.01/0.01)

: negative (12.32/2.0)

Number of Rules : 28

Time taken to build model: 0.21 seconds

==== Evaluation on test split ====

Time taken to test model on test split: 0 seconds

==== Summary ====

Correctly Classified Instances      1255          97.8939 %
Incorrectly Classified Instances   27           2.1061 %
Kappa statistic                   0.7974
Mean absolute error               0.0256
Root mean squared error          0.1302
Relative absolute error           22.4569 %
Root relative squared error     55.4605 %
Total Number of Instances        1282

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.993     0.240     0.985     0.993     0.989     0.799     0.963     0.995    negative
      0.760     0.007     0.864     0.760     0.809     0.799     0.963     0.795    sick
Weighted Avg.      0.979     0.226     0.978     0.979     0.978     0.799     0.963     0.983

==== Confusion Matrix ====

      a      b  <-- classified as
1198    9 |    a = negative
      18    57 |    b = sick

```

## Leave one out fold

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B10-M-1.0-Rfirst-last-precision6
Instances:   3772

```

```

Attributes:    78
age_1
age_2
age_3
age_4
age_5
age_6
age_7
age_8
age_9
sex
on_thyroxine
query_on_thyroxine
on_antithyroid_medication
sick
pregnant
thyroid_surgery
I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH_1
TSH_2
TSH_3
TSH_4
TSH_5
TSH_6
TSH_7
TSH_8
TSH_9
T3_measured
T3_1
T3_2
T3_3
T3_4
T3_5
T3_6
T3_7
T3_8
T3_9
TT4_measured
TT4_1
TT4_2
TT4_3
TT4_4
TT4_5
TT4_6
TT4_7
TT4_8
TT4_9
T4U_measured
T4U_1
T4U_2
T4U_3
T4U_4
T4U_5
T4U_6
T4U_7
T4U_8
T4U_9
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
FTI_5
FTI_6
FTI_7
FTI_8
FTI_9
TBG_measured
TBG
referral_source
Class
Test mode: 3772-fold cross-validation

==== Classifier model (full training set) ===

PART decision list
-----
T3_1 = '(1.105-inf)' AND

```

```

FTI_4 = '(-inf-159.2)': negative (3170.02/7.69)
referral_source = other AND
T3_measured = f: negative (118.6/0.1)

T3_1 = '(1.105-inf)' AND
T3_3 = '(3.215-inf)': negative (60.01/0.03)

T3_1 = '(1.105-inf)' AND
on_thyroxine = t: negative (34.86)

TT4_1 = '(-inf-44.8]' AND
TSH_1 = '(53.0045-inf)': negative (26.16)

T3_measured = t AND
T3_1 = '(1.105-inf)' AND
TT4_3 = '(-inf-130.4)': negative (13.44)

T3_measured = t AND
FTI_4 = '(-inf-159.2)' AND
TT4_1 = '(44.8-inf)' AND
referral_source = SVI AND
FTI_2 = '(80.6-inf)': sick (146.27/2.03)

sick = f AND
T3_measured = t AND
TT4_5 = '(-inf-216)' AND
T4U_1 = '(0.457-inf)' AND
FTI_1 = '(41.3-inf)' AND
T4U_4 = '(-inf-1.078]' AND
FTI_4 = '(159.2-inf)' AND
query_hypothyroid = f AND
TSH_measured = t AND
on_thyroxine = f AND
query_hyperthyroid = f AND
TT4_2 = '(87.6-inf)' AND
referral_source = SVI AND
T3_2 = '(-inf-2.16]': negative (23.79/4.79)

sick = f AND
T3_measured = t AND
TT4_5 = '(-inf-216)' AND
T4U_1 = '(0.457-inf)' AND
FTI_1 = '(-inf-41.3)': negative (12.58/1.13)

T3_measured = f: negative (10.18/0.18)

sick = t: sick (8.34)

TT4_5 = '(-inf-216)' AND
T4U_1 = '(0.457-inf)' AND
T4U_4 = '(-inf-1.078]' AND
TSH_measured = t AND
referral_source = SVHC AND
T3_1 = '(-inf-1.105)': sick (8.0)

psych = f AND
TT4_5 = '(-inf-216)' AND
T4U_1 = '(0.457-inf)' AND
T4U_4 = '(-inf-1.078]' AND
referral_source = other AND
T4U_measured = t AND
T3_2 = '(2.16-inf)': negative (9.0/2.0)

psych = f AND
TT4_5 = '(-inf-216)' AND
T4U_1 = '(-inf-0.457)': negative (5.28/0.06)

psych = f AND
TT4_5 = '(-inf-216)' AND
T4U_4 = '(1.078-inf)': negative (5.28/0.06)

psych = f AND
TT4_5 = '(-inf-216)' AND
referral_source = other AND
T4U_measured = t AND
TT4_4 = '(-inf-173.2)' AND
query_hypothyroid = t AND
FTI_2 = '(80.6-inf)': sick (4.0)

psych = f AND
TT4_5 = '(-inf-216)' AND
referral_source = other AND
T4U_measured = t AND
query_hypothyroid = f AND
TT4_4 = '(-inf-173.2)' AND
TSH_2 = '(-inf-106.004]' AND

```

```

T4U_2 = '(0.664-inf)' AND
T4U_3 = '(-inf-0.871]' AND
age_1 = '(-inf-46.4]': negative (5.96)

psych = f AND
TT4_5 = '(-inf-216]' AND
referral_source = other AND
T4U_measured = t AND
query_hypothyroid = f AND
TT4_4 = '(-inf-173.2]' AND
TSH_2 = '(-inf-106.004]' AND
T4U_2 = '(0.664-inf)' AND
sex = M: negative (11.47/3.26)

psych = f AND
TT4_5 = '(-inf-216]' AND
query_hypothyroid = t AND
T4U_3 = '(-inf-0.871]': negative (5.0)

psych = f AND
TT4_5 = '(-inf-216]' AND
referral_source = SVI AND
sex = F AND
age_1 = '(46.4-inf)' AND
T3_1 = '(-inf-1.105]': negative (18.67/6.83)

referral_source = SVI AND
TT4_4 = '(-inf-173.2]' AND
T3_1 = '(-inf-1.105]': sick (15.54/1.0)

T4U_measured = t AND
psych = f AND
TT4_5 = '(-inf-216]' AND
referral_source = SVI AND
age_1 = '(46.4-inf)': sick (5.0/1.0)

T3_2 = '(2.16-inf)': negative (7.65)

T4U_measured = t AND
psych = f AND
referral_source = other AND
query_hypothyroid = f AND
TT4_4 = '(-inf-173.2]' AND
TSH_2 = '(-inf-106.004]' AND
T4U_2 = '(0.664-inf)' AND
on_thyroxine = f AND
age_1 = '(-inf-46.4]' AND
T3_1 = '(-inf-1.105]': negative (4.73)

T4U_measured = t AND
psych = f AND
referral_source = other AND
query_hypothyroid = f AND
age_1 = '(46.4-inf)' AND
TT4_4 = '(-inf-173.2]' AND
TSH_1 = '(-inf-53.0045]' AND
FTI_4 = '(-inf-159.2]' AND
T4U_2 = '(0.664-inf)' AND
on_thyroxine = f: sick (10.09/3.05)

T4U_measured = t AND
psych = f AND
referral_source = other AND
query_hypothyroid = f AND
age_1 = '(46.4-inf)' AND
TT4_2 = '(87.6-inf)' AND
TT4_4 = '(-inf-173.2]': negative (8.74/1.69)

T4U_measured = t AND
psych = f AND
referral_source = other AND
query_hypothyroid = f: sick (11.01/0.01)

: negative (12.32/2.0)

Number of Rules : 28

Time taken to build model: 0.23 seconds

==== Stratified cross-validation ===
==== Summary ===

Correctly Classified Instances      3692          97.8791 %
Incorrectly Classified Instances     80           2.1209 %
Kappa statistic                      0.8117
Mean absolute error                  0.026

```

```

Root mean squared error          0.1293
Relative absolute error         22.562 %
Root relative squared error    53.908 %
Total Number of Instances      3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC     ROC Area   PRC Area   Class
      0.990     0.195     0.987     0.990     0.989     0.812     0.941     0.989   negative
      0.805     0.010     0.842     0.805     0.823     0.812     0.941     0.852   sick
Weighted Avg.      0.979     0.183     0.978     0.979     0.979     0.812     0.941     0.980

==== Confusion Matrix ====

      a     b  <-- classified as
3506  35 |     a = negative
      45  186 |     b = sick

```

### 3.1.4 40 bins

#### Cross Validation 5 folds

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B40-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  258
              [list of attributes omitted]
Test mode:   5-fold cross-validation

==== Classifier model (full training set) ====

PART decision list
-----
T3_4 = '(1.105-inf)' AND
FTI_17 = '(-inf-169.025]' AND
TT4_5 = '(55.5-inf)': negative (3174.48/2.9)

referral_source = other AND
sick = f AND
age_6 = '(-inf-69.1]' AND
I131_treatment = f AND
TSH_measured = t AND
T4U_15 = '(1.02625-inf)': negative (69.64)

T3_measured = f: negative (106.58/2.1)

T3_10 = '(2.6875-inf)': negative (55.39)

TT4_5 = '(-inf-55.5]' AND
TSH_2 = '(26.50475-inf)': negative (37.13)

TT4_18 = '(-inf-194.6]' AND
referral_source = SVI AND
TSH_1 = '(-inf-13.254875]' AND
FTI_16 = '(-inf-159.2]' AND
TT4_5 = '(55.5-inf)' AND
TT4_6 = '(60.95-inf)': sick (156.47/0.58)

sick = f AND
TT4_18 = '(-inf-194.6]' AND
age_1 = '(12.35-inf)' AND
T4U_measured = t AND
T4U_4 = '(0.457-inf)' AND
T4U_5 = '(0.50875-inf)' AND
FTI_1 = '(11.825-inf)' AND
T4U_17 = '(-inf-1.12975]' AND
FTI_4 = '(41.3-inf)' AND
TSH_2 = '(-inf-26.50475]' AND
T3_9 = '(-inf-2.42375]' AND
referral_source = other AND
T3_2 = '(0.5775-inf)' AND
query_hypothyroid = f AND
T4U_11 = '(-inf-0.81925]': negative (16.9)

sick = f AND
TT4_18 = '(-inf-194.6]' AND
age_1 = '(12.35-inf)' AND
T4U_measured = f: negative (12.02/0.93)

sick = f AND

```

```

TT4_18 = '(-inf-194.6]' AND
T4U_4 = '(0.457-inf)' AND
TT4_1 = '(12.7-inf)' AND
age_1 = '(12.35-inf)' AND
T4U_17 = '(-inf-1.12975]' AND
FTI_4 = '(41.3-inf)' AND
T4U_5 = '(0.50875-inf)' AND
TSH_2 = '(-inf-26.50475]' AND
T3_9 = '(-inf-2.42375]' AND
age_3 = '(-inf-35.05]' AND
T3_8 = '(-inf-2.16]': sick (9.0/1.0)

sick = f AND
TT4_18 = '(-inf-194.6]' AND
age_2 = '(23.7-inf)' AND
FTI_1 = '(11.825-inf)' AND
T4U_3 = '(0.40525-inf)' AND
T4U_5 = '(0.50875-inf)' AND
T4U_17 = '(-inf-1.12975]' AND
FTI_4 = '(41.3-inf)' AND
TSH_2 = '(-inf-26.50475]' AND
age_3 = '(35.05-inf)' AND
T3_8 = '(-inf-2.16]' AND
FTI_15 = '(149.375-inf)' AND
TT4_16 = '(-inf-173.2]': negative (24.04/3.0)

TT4_18 = '(194.6-inf)': negative (8.0)

age_2 = '(23.7-inf)' AND
TT4_1 = '(12.7-inf)' AND
T4U_3 = '(0.40525-inf)' AND
T4U_17 = '(-inf-1.12975]' AND
age_7 = '(-inf-80.45]' AND
sick = f AND
FTI_20 = '(-inf-198.5]' AND
FTI_4 = '(41.3-inf)' AND
T4U_5 = '(0.50875-inf)' AND
TSH_2 = '(-inf-26.50475]' AND
age_3 = '(35.05-inf)' AND
T3_8 = '(-inf-2.16]' AND
FTI_19 = '(-inf-188.675]' AND
FTI_13 = '(-inf-129.725]' AND
T3_6 = '(-inf-1.6325]' AND
psych = f AND
referral_source = other AND
age_5 = '(-inf-57.75]': negative (7.03)

age_2 = '(23.7-inf)' AND
TT4_1 = '(12.7-inf)' AND
T4U_3 = '(0.40525-inf)' AND
T4U_17 = '(-inf-1.12975]' AND
TT4_6 = '(66.2-inf)' AND
age_3 = '(35.05-inf)': sick (42.24/4.07)

FTI_14 = '(-inf-139.55]' AND
age_7 = '(-inf-80.45]' AND
FTI_12 = '(-inf-119.9]' AND
TT4_1 = '(12.7-inf)' AND
query_hypothyroid = f AND
on_thyroxine = f AND
FTI_8 = '(-inf-80.6]' AND
query_hyperthyroid = f AND
FTI_4 = '(41.3-inf)' AND
T3_6 = '(-inf-1.6325]' AND
age_4 = '(46.4-inf)' AND
sex = F AND
T4U_6 = '(0.5605-inf)': negative (9.0/1.0)

FTI_14 = '(-inf-139.55]' AND
query_hypothyroid = f AND
FTI_1 = '(11.825-inf)' AND
age_2 = '(23.7-inf)' AND
query_hyperthyroid = f AND
T3_6 = '(-inf-1.6325]' AND
T4U_8 = '(0.664-inf)': sick (15.0/1.0)

age_4 = '(-inf-46.4]': negative (10.05/0.01)

TT4_5 = '(-inf-55.5]' AND
TT4_4 = '(44.8-inf)': negative (8.01)

FTI_12 = '(-inf-119.9]' AND
referral_source = SVI AND
age_5 = '(57.75-inf)': negative (4.0/1.0)

TSH_1 = '(-inf-13.254875]': sick (5.01/1.01)

```

```

: negative (2.0)

Number of Rules : 20

Time taken to build model: 0.67 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3709          98.3298 %
Incorrectly Classified Instances   63           1.6702 %
Kappa statistic                   0.8527
Mean absolute error               0.0204
Root mean squared error          0.1246
Relative absolute error           17.6862 %
Root relative squared error     51.9465 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC    ROC Area  PRC Area  Class
0.992    0.152    0.990     0.992    0.991     0.853    0.933    0.986    negative
0.848    0.008    0.875     0.848    0.862     0.853    0.933    0.862    sick
Weighted Avg.       0.983    0.143    0.983     0.983    0.983     0.853    0.933    0.978

==== Confusion Matrix ====

      a      b  <-- classified as
3513   28 |      a = negative
      35   196 |      b = sick

```

## Cross Validation 10 folds

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B40-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  258
              [list of attributes omitted]
Test mode:   10-fold cross-validation

==== Classifier model (full training set) ====

PART decision list
-----
T3_4 = '(1.105-inf)' AND
FTI_17 = '(-inf-169.025]' AND
TT4_5 = '(55.5-inf)': negative (3174.48/2.9)

referral_source = other AND
sick = f AND
age_6 = '(-inf-69.1]' AND
I131_treatment = f AND
TSH_measured = t AND
T4U_15 = '(1.02625-inf)': negative (69.64)

T3_measured = f: negative (106.58/2.1)

T3_10 = '(2.6875-inf)': negative (55.39)

TT4_5 = '(-inf-55.5]' AND
TSH_2 = '(26.50475-inf)': negative (37.13)

TT4_18 = '(-inf-194.6]' AND
referral_source = SVI AND
TSH_1 = '(-inf-13.254875]' AND
FTI_16 = '(-inf-159.2]' AND
TT4_5 = '(55.5-inf)' AND
TT4_6 = '(60.95-inf)': sick (156.47/0.58)

sick = f AND
TT4_18 = '(-inf-194.6]' AND
age_1 = '(12.35-inf)' AND
T4U_measured = t AND
T4U_4 = '(0.457-inf)' AND
T4U_5 = '(0.50875-inf)' AND
FTI_1 = '(11.825-inf)' AND
T4U_17 = '(-inf-1.12975]' AND
FTI_4 = '(41.3-inf)' AND
TSH_2 = '(-inf-26.50475]' AND
T3_9 = '(-inf-2.42375]' AND

```

```

referral_source = other AND
T3_2 = '(0.5775-inf)' AND
query_hypothyroid = f AND
T4U_11 = '(-inf-0.81925]': negative (16.9)

sick = f AND
TT4_18 = '(-inf-194.6]' AND
age_1 = '(12.35-inf)' AND
T4U_measured = f: negative (12.02/0.93)

sick = f AND
TT4_18 = '(-inf-194.6]' AND
T4U_4 = '(0.457-inf)' AND
TT4_1 = '(12.7-inf)' AND
age_1 = '(12.35-inf)' AND
T4U_17 = '(-inf-1.12975]' AND
FTI_4 = '(41.3-inf)' AND
T4U_5 = '(0.50875-inf)' AND
TSH_2 = '(-inf-26.50475]' AND
T3_9 = '(-inf-2.42375]' AND
age_3 = '(-inf-35.05)' AND
T3_8 = '(-inf-2.16)': sick (9.0/1.0)

sick = f AND
TT4_18 = '(-inf-194.6]' AND
age_2 = '(23.7-inf)' AND
FTI_1 = '(11.825-inf)' AND
T4U_3 = '(0.40525-inf)' AND
T4U_5 = '(0.50875-inf)' AND
T4U_17 = '(-inf-1.12975]' AND
FTI_4 = '(41.3-inf)' AND
TSH_2 = '(-inf-26.50475)' AND
age_3 = '(35.05-inf)' AND
T3_8 = '(-inf-2.16)' AND
FTI_15 = '(149.375-inf)' AND
FTI_16 = '(-inf-173.2)': negative (24.04/3.0)

TT4_18 = '(194.6-inf)': negative (8.0)

age_2 = '(23.7-inf)' AND
TT4_1 = '(12.7-inf)' AND
T4U_3 = '(0.40525-inf)' AND
T4U_17 = '(-inf-1.12975]' AND
age_7 = '(-inf-80.45)' AND
sick = f AND
FTI_20 = '(-inf-198.5)' AND
FTI_4 = '(41.3-inf)' AND
T4U_5 = '(0.50875-inf)' AND
TSH_2 = '(-inf-26.50475]' AND
age_3 = '(35.05-inf)' AND
T3_8 = '(-inf-2.16)' AND
T3_19 = '(-inf-188.675)' AND
FTI_13 = '(-inf-129.725)' AND
T3_6 = '(-inf-1.6325)' AND
psych = f AND
referral_source = other AND
age_5 = '(-inf-57.75)': negative (7.03)

age_2 = '(23.7-inf)' AND
TT4_1 = '(12.7-inf)' AND
T4U_3 = '(0.40525-inf)' AND
T4U_17 = '(-inf-1.12975]' AND
TT4_6 = '(66.2-inf)' AND
age_3 = '(35.05-inf)': sick (42.24/4.07)

FTI_14 = '(-inf-139.55)' AND
age_7 = '(-inf-80.45)' AND
FTI_12 = '(-inf-119.9)' AND
TT4_1 = '(12.7-inf)' AND
query_hypothyroid = f AND
on_thyroxine = f AND
FTI_8 = '(-inf-80.6)' AND
query_hyperthyroid = f AND
FTI_4 = '(41.3-inf)' AND
T3_6 = '(-inf-1.6325)' AND
age_4 = '(46.4-inf)' AND
sex = F AND
T4U_6 = '(0.5605-inf)': negative (9.0/1.0)

FTI_14 = '(-inf-139.55)' AND
query_hypothyroid = f AND
FTI_1 = '(11.825-inf)' AND
age_2 = '(23.7-inf)' AND
query_hyperthyroid = f AND
T3_6 = '(-inf-1.6325)' AND
T4U_8 = '(0.664-inf)': sick (15.0/1.0)

```

```

age_4 = '(-inf-46.4]': negative (10.05/0.01)
TT4_5 = '(-inf-55.5]' AND
TT4_4 = '(44.8-inf)': negative (8.01)

FTI_12 = '(-inf-119.9]' AND
referral_source = SVI AND
age_5 = '(57.75-inf)': negative (4.0/1.0)

TSH_1 = '(-inf-13.254875]': sick (5.01/1.01)

: negative (2.0)

Number of Rules : 20

Time taken to build model: 0.64 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3715           98.4889 %
Incorrectly Classified Instances   57            1.5111 %
Kappa statistic                   0.8683
Mean absolute error               0.0183
Root mean squared error          0.117
Relative absolute error           15.8926 %
Root relative squared error     48.8156 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.992     0.126     0.992     0.992     0.992     0.868     0.934     0.986     negative
      0.874     0.008     0.878     0.874     0.876     0.868     0.934     0.857     sick
Weighted Avg.      0.985     0.118     0.985     0.985     0.985     0.868     0.934     0.978

==== Confusion Matrix ====

      a      b  <-- classified as
3513  28 |    a = negative
      29  202 |    b = sick

```

## Cross Validation 15 folds

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B40-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  258
              [list of attributes omitted]
Test mode:   15-fold cross-validation

==== Classifier model (full training set) ====

PART decision list
-----
T3_4 = '(1.105-inf)' AND
FTI_17 = '(-inf-169.025]' AND
TT4_5 = '(55.5-inf)': negative (3174.48/2.9)

referral_source = other AND
sick = f AND
age_6 = '(-inf-69.1]' AND
T131_treatment = f AND
TSH_measured = t AND
T4U_15 = '(1.02625-inf)': negative (69.64)

T3_measured = f: negative (106.58/2.1)

T3_10 = '(2.6875-inf)': negative (55.39)

TT4_5 = '(-inf-55.5]' AND
TSH_2 = '(26.50475-inf)': negative (37.13)

TT4_18 = '(-inf-194.6]' AND
referral_source = SVI AND
TSH_1 = '(-inf-13.254875]': sick (156.47/0.58)
FTI_16 = '(-inf-159.2]' AND
TT4_5 = '(55.5-inf)' AND
FTI_6 = '(60.95-inf)': sick (156.47/0.58)

```

```

sick = f AND
TT4_18 = '(-inf-194.6)' AND
age_1 = '(12.35-inf)' AND
T4U_measured = t AND
T4U_4 = '(0.457-inf)' AND
T4U_5 = '(0.50875-inf)' AND
FTI_1 = '(11.825-inf)' AND
T4U_17 = '(-inf-1.12975]' AND
FTI_4 = '(41.3-inf)' AND
TSH_2 = '(-inf-26.50475]' AND
T3_9 = '(-inf-2.42375)' AND
referral_source = other AND
T3_2 = '(0.5775-inf)' AND
query_hypothyroid = f AND
T4U_11 = '(-inf-0.81925)': negative (16.9)

sick = f AND
TT4_18 = '(-inf-194.6)' AND
age_1 = '(12.35-inf)' AND
T4U_measured = f: negative (12.02/0.93)

sick = f AND
TT4_18 = '(-inf-194.6)' AND
T4U_4 = '(0.457-inf)' AND
TT4_1 = '(12.7-inf)' AND
age_1 = '(12.35-inf)' AND
T4U_17 = '(-inf-1.12975]' AND
FTI_4 = '(41.3-inf)' AND
T4U_5 = '(0.50875-inf)' AND
TSH_2 = '(-inf-26.50475]' AND
T3_9 = '(-inf-2.42375)' AND
age_3 = '(-inf-35.05)' AND
T3_8 = '(-inf-2.16)': sick (9.0/1.0)

sick = f AND
TT4_18 = '(-inf-194.6)' AND
age_2 = '(23.7-inf)' AND
FTI_1 = '(11.825-inf)' AND
T4U_3 = '(0.40525-inf)' AND
T4U_5 = '(0.50875-inf)' AND
T4U_17 = '(-inf-1.12975]' AND
FTI_4 = '(41.3-inf)' AND
TSH_2 = '(-inf-26.50475]' AND
age_3 = '(35.05-inf)' AND
T3_8 = '(-inf-2.16)' AND
FTI_15 = '(149.375-inf)' AND
TT4_16 = '(-inf-173.2)': negative (24.04/3.0)

TT4_18 = '(194.6-inf)': negative (8.0)

age_2 = '(23.7-inf)' AND
TT4_1 = '(12.7-inf)' AND
T4U_3 = '(0.40525-inf)' AND
T4U_17 = '(-inf-1.12975]' AND
age_7 = '(-inf-80.45)' AND
sick = f AND
FTI_20 = '(-inf-198.5)' AND
FTI_4 = '(41.3-inf)' AND
T4U_5 = '(0.50875-inf)' AND
TSH_2 = '(-inf-26.50475)' AND
age_3 = '(35.05-inf)' AND
T3_8 = '(-inf-2.16)' AND
FTI_19 = '(-inf-188.675)' AND
FTI_13 = '(-inf-129.725)' AND
T3_6 = '(-inf-1.6325)' AND
psych = f AND
referral_source = other AND
age_5 = '(-inf-57.75)': negative (7.03)

age_2 = '(23.7-inf)' AND
TT4_1 = '(12.7-inf)' AND
T4U_3 = '(0.40525-inf)' AND
T4U_17 = '(-inf-1.12975]' AND
TT4_6 = '(66.2-inf)' AND
age_3 = '(35.05-inf)': sick (42.24/4.07)

FTI_14 = '(-inf-139.55)' AND
age_7 = '(-inf-80.45)' AND
FTI_12 = '(-inf-119.9)' AND
TT4_1 = '(12.7-inf)' AND
query_hypothyroid = f AND
on_thyroxine = f AND
FTI_8 = '(-inf-80.6)' AND
query_hyperthyroid = f AND
FTI_4 = '(41.3-inf)' AND
T3_6 = '(-inf-1.6325)' AND
age_4 = '(46.4-inf)' AND

```

```

sex = F AND
T4U_6 = '(0.5605-inf)': negative (9.0/1.0)

FTI_14 = '(-inf-139.55]' AND
query_hypothyroid = f AND
FTI_1 = '(11.825-inf)' AND
age_2 = '(23.7-inf)' AND
query_hyperthyroid = f AND
T3_6 = '(-inf-1.6325]' AND
T4U_8 = '(0.664-inf)': sick (15.0/1.0)

age_4 = '(-inf-46.4]': negative (10.05/0.01)

TT4_5 = '(-inf-55.5]' AND
TT4_4 = '(44.8-inf)': negative (8.01)

FTI_12 = '(-inf-119.9]' AND
referral_source = SVI AND
age_5 = '(57.75-inf)': negative (4.0/1.0)

TSH_1 = '(-inf-13.254875)': sick (5.01/1.01)

: negative (2.0)

Number of Rules : 20

Time taken to build model: 0.69 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3712          98.4093 %
Incorrectly Classified Instances    60           1.5907 %
Kappa statistic                   0.8611
Mean absolute error               0.0185
Root mean squared error          0.1199
Relative absolute error          16.0282 %
Root relative squared error     50.0036 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.992     0.134     0.991     0.992     0.992     0.861     0.929     0.984     negative
      0.866     0.008     0.873     0.866     0.870     0.861     0.929     0.867     sick
Weighted Avg.                      0.984     0.126     0.984     0.984     0.984     0.861     0.929     0.977

==== Confusion Matrix ====

      a      b  <-- classified as
3512   29 |      a = negative
      31   200 |      b = sick

```

## Percentage Split 50%

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B40-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  258
              [list of attributes omitted]
Test mode:   split 50.0% train, remainder test

==== Classifier model (full training set) ====

PART decision list
-----
T3_4 = '(1.105-inf)' AND
FTI_17 = '(-inf-169.025]' AND
TT4_5 = '(55.5-inf)': negative (3174.48/2.9)

referral_source = other AND
sick = f AND
age_6 = '(-inf-69.1]' AND
I131_treatment = f AND
TSH_measured = t AND
T4U_15 = '(1.02625-inf)': negative (69.64)

T3_measured = f: negative (106.58/2.1)

T3_10 = '(2.6875-inf)': negative (55.39)

```

```

TT4_5 = '(-inf-55.5]' AND
TSH_2 = '(26.50475-inf)': negative (37.13)

TT4_18 = '(-inf-194.6]' AND
referral_source = SVI AND
TSH_1 = '(-inf-13.254875]' AND
FTI_16 = '(-inf-159.2]' AND
TT4_5 = '(55.5-inf)' AND
FTI_6 = '(60.95-inf)': sick (156.47/0.58)

sick = f AND
TT4_18 = '(-inf-194.6]' AND
age_1 = '(12.35-inf)' AND
T4U_measured = t AND
T4U_4 = '(0.457-inf)' AND
T4U_5 = '(0.50875-inf)' AND
FTI_1 = '(11.825-inf)' AND
T4U_17 = '(-inf-1.12975]' AND
FTI_4 = '(41.3-inf)' AND
TSH_2 = '(-inf-26.50475]' AND
T3_9 = '(-inf-2.42375]' AND
referral_source = other AND
T3_2 = '(0.5775-inf)' AND
query_hypothyroid = f AND
T4U_11 = '(-inf-0.81925)': negative (16.9)

sick = f AND
TT4_18 = '(-inf-194.6]' AND
age_1 = '(12.35-inf)' AND
T4U_measured = f: negative (12.02/0.93)

sick = f AND
TT4_18 = '(-inf-194.6]' AND
T4U_4 = '(0.457-inf)' AND
TT4_1 = '(12.7-inf)' AND
age_1 = '(12.35-inf)' AND
T4U_17 = '(-inf-1.12975]' AND
FTI_4 = '(41.3-inf)' AND
T4U_5 = '(0.50875-inf)' AND
TSH_2 = '(-inf-26.50475]' AND
T3_9 = '(-inf-2.42375]' AND
age_3 = '(-inf-35.05]' AND
T3_8 = '(-inf-2.16)': sick (9.0/1.0)

sick = f AND
TT4_18 = '(-inf-194.6]' AND
age_2 = '(23.7-inf)' AND
FTI_1 = '(11.825-inf)' AND
T4U_3 = '(0.40525-inf)' AND
T4U_5 = '(0.50875-inf)' AND
T4U_17 = '(-inf-1.12975]' AND
FTI_4 = '(41.3-inf)' AND
TSH_2 = '(-inf-26.50475]' AND
age_3 = '(35.05-inf)' AND
T3_8 = '(-inf-2.16)' AND
FTI_15 = '(149.375-inf)' AND
TT4_16 = '(-inf-173.2)': negative (24.04/3.0)

TT4_18 = '(194.6-inf)': negative (8.0)

age_2 = '(23.7-inf)' AND
TT4_1 = '(12.7-inf)' AND
T4U_3 = '(0.40525-inf)' AND
T4U_17 = '(-inf-1.12975]' AND
age_7 = '(-inf-80.45]' AND
sick = f AND
FTI_20 = '(-inf-198.5]' AND
FTI_4 = '(41.3-inf)' AND
T4U_5 = '(0.50875-inf)' AND
TSH_2 = '(-inf-26.50475]' AND
age_3 = '(35.05-inf)' AND
T3_8 = '(-inf-2.16)' AND
FTI_19 = '(-inf-188.675]' AND
FTI_13 = '(-inf-129.725]' AND
T3_6 = '(-inf-1.6325]' AND
psych = f AND
referral_source = other AND
age_5 = '(-inf-57.75)': negative (7.03)

age_2 = '(23.7-inf)' AND
TT4_1 = '(12.7-inf)' AND
T4U_3 = '(0.40525-inf)' AND
T4U_17 = '(-inf-1.12975]' AND
TT4_6 = '(66.2-inf)' AND
age_3 = '(35.05-inf)': sick (42.24/4.07)

```

```

FTI_14 = '(-inf-139.55]' AND
age_7 = '(-inf-80.45]' AND
FTI_12 = '(-inf-119.9]' AND
TT4_1 = '(12.7-inf)' AND
query_hypothyroid = f AND
on_thyroxine = f AND
FTI_8 = '(-inf-80.6]' AND
query_hyperthyroid = f AND
FTI_4 = '(41.3-inf)' AND
T3_6 = '(-inf-1.6325]' AND
age_4 = '(46.4-inf)' AND
sex = F AND
T4U_6 = '(0.5605-inf)': negative (9.0/1.0)

FTI_14 = '(-inf-139.55]' AND
query_hypothyroid = f AND
FTI_1 = '(11.825-inf)' AND
age_2 = '(23.7-inf)' AND
query_hyperthyroid = f AND
T3_6 = '(-inf-1.6325]' AND
T4U_8 = '(0.664-inf)': sick (15.0/1.0)

age_4 = '(-inf-46.4)': negative (10.05/0.01)

TT4_5 = '(-inf-55.5]' AND
TT4_4 = '(44.8-inf)': negative (8.01)

FTI_12 = '(-inf-119.9]' AND
referral_source = SVI AND
age_5 = '(57.75-inf)': negative (4.0/1.0)

TSH_1 = '(-inf-13.254875)': sick (5.01/1.01)

: negative (2.0)

Number of Rules : 20

Time taken to build model: 0.65 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0.02 seconds

==== Summary ===

    Correctly Classified Instances      1844           97.7731 %
    Incorrectly Classified Instances     42            2.2269 %
    Kappa statistic                      0.8007
    Mean absolute error                  0.0259
    Root mean squared error              0.1423
    Relative absolute error              22.4692 %
    Root relative squared error          59.0012 %
    Total Number of Instances           1886

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.991     0.222     0.985      0.991     0.988      0.802     0.969     0.997     negative
      0.778     0.009     0.850      0.778     0.813      0.802     0.969     0.797     sick
Weighted Avg.   0.978     0.209     0.977      0.978     0.977      0.802     0.969     0.984

==== Confusion Matrix ===

      a      b  <-- classified as
1753   16 |   a = negative
      26   91 |   b = sick

```

## Percentage Split 60%

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B40-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  258
              [list of attributes omitted]
Test mode:   split 60.0% train, remainder test

==== Classifier model (full training set) ===

PART decision list
-----

```

```

T3_4 = '(1.105-inf)' AND
FTI_17 = '(-inf-169.025]' AND
TT4_5 = '(55.5-inf)': negative (3174.48/2.9)

referral_source = other AND
sick = f AND
age_6 = '(-inf-69.1]' AND
I131_treatment = f AND
TSH_measured = t AND
T4U_15 = '(1.02625-inf)': negative (69.64)

T3_measured = f: negative (106.58/2.1)

T3_10 = '(2.6875-inf)': negative (55.39)

TT4_5 = '(-inf-55.5]' AND
TSH_2 = '(26.50475-inf)': negative (37.13)

TT4_18 = '(-inf-194.6]' AND
referral_source = SVI AND
TSH_1 = '(-inf-13.254875]' AND
FTI_16 = '(-inf-159.2]' AND
TT4_5 = '(55.5-inf)' AND
FTI_6 = '(60.95-inf)': sick (156.47/0.58)

sick = f AND
TT4_18 = '(-inf-194.6]' AND
age_1 = '(12.35-inf)' AND
T4U_measured = t AND
T4U_4 = '(0.457-inf)' AND
T4U_5 = '(0.50875-inf)' AND
FTI_1 = '(11.825-inf)' AND
T4U_17 = '(-inf-1.12975]' AND
FTI_4 = '(41.3-inf)' AND
TSH_2 = '(-inf-26.50475]' AND
T3_9 = '(-inf-2.42375]' AND
referral_source = other AND
T3_2 = '(0.5775-inf)' AND
query_hypothyroid = f AND
T4U_11 = '(-inf-0.81925)': negative (16.9)

sick = f AND
TT4_18 = '(-inf-194.6]' AND
age_1 = '(12.35-inf)' AND
T4U_measured = f: negative (12.02/0.93)

sick = f AND
TT4_18 = '(-inf-194.6]' AND
T4U_4 = '(0.457-inf)' AND
TT4_1 = '(12.7-inf)' AND
age_1 = '(12.35-inf)' AND
T4U_17 = '(-inf-1.12975]' AND
FTI_4 = '(41.3-inf)' AND
T4U_5 = '(0.50875-inf)' AND
TSH_2 = '(-inf-26.50475]' AND
T3_9 = '(-inf-2.42375]' AND
age_3 = '(-inf-35.05)' AND
T3_8 = '(-inf-2.16)': sick (9.0/1.0)

sick = f AND
TT4_18 = '(-inf-194.6]' AND
age_2 = '(23.7-inf)' AND
FTI_1 = '(11.825-inf)' AND
T4U_3 = '(0.40525-inf)' AND
T4U_5 = '(0.50875-inf)' AND
T4U_17 = '(-inf-1.12975]' AND
FTI_4 = '(41.3-inf)' AND
TSH_2 = '(-inf-26.50475]' AND
age_3 = '(35.05-inf)' AND
T3_8 = '(-inf-2.16)' AND
FTI_15 = '(149.375-inf)' AND
TT4_16 = '(-inf-173.2)': negative (24.04/3.0)

TT4_18 = '(194.6-inf)': negative (8.0)

age_2 = '(23.7-inf)' AND
TT4_1 = '(12.7-inf)' AND
T4U_3 = '(0.40525-inf)' AND
T4U_17 = '(-inf-1.12975]' AND
age_7 = '(-inf-80.45)' AND
sick = f AND
FTI_20 = '(-inf-198.5)' AND
FTI_4 = '(41.3-inf)' AND
T4U_5 = '(0.50875-inf)' AND
TSH_2 = '(-inf-26.50475]' AND
age_3 = '(35.05-inf)' AND
T3_8 = '(-inf-2.16)' AND

```

```

FTI_19 = '(-inf-188.675)' AND
FTI_13 = '(-inf-129.725)' AND
T3_6 = '(-inf-1.6325)' AND
psych = f AND
referral_source = other AND
age_5 = '(-inf-57.75)': negative (7.03)

age_2 = '(23.7-inf)' AND
TT4_1 = '(12.7-inf)' AND
T4U_3 = '(0.40525-inf)' AND
T4U_17 = '(-inf-1.12975)' AND
TT4_6 = '(66.2-inf)' AND
age_3 = '(35.05-inf)': sick (42.24/4.07)

FTI_14 = '(-inf-139.55)' AND
age_7 = '(-inf-80.45)' AND
FTI_12 = '(-inf-119.9)' AND
TT4_1 = '(12.7-inf)' AND
query_hypothyroid = f AND
on_thyroxine = f AND
FTI_8 = '(-inf-80.6]' AND
query_hyperthyroid = f AND
FTI_4 = '(41.3-inf)' AND
T3_6 = '(-inf-1.6325)' AND
age_4 = '(46.4-inf)' AND
sex = F AND
T4U_6 = '(0.5605-inf)': negative (9.0/1.0)

FTI_14 = '(-inf-139.55)' AND
query_hypothyroid = f AND
FTI_1 = '(11.825-inf)' AND
age_2 = '(23.7-inf)' AND
query_hyperthyroid = f AND
T3_6 = '(-inf-1.6325)' AND
T4U_8 = '(0.664-inf)': sick (15.0/1.0)

age_4 = '(-inf-46.4)': negative (10.05/0.01)

TT4_5 = '(-inf-55.5)' AND
TT4_4 = '(44.8-inf)': negative (8.01)

FTI_12 = '(-inf-119.9)' AND
referral_source = SVI AND
age_5 = '(57.75-inf)': negative (4.0/1.0)

TSH_1 = '(-inf-13.254875)': sick (5.01/1.01)

: negative (2.0)

Number of Rules : 20

Time taken to build model: 0.77 seconds

==== Evaluation on test split ====

Time taken to test model on test split: 0.04 seconds

==== Summary ====

Correctly Classified Instances      1481          98.1445 %
Incorrectly Classified Instances    28           1.8555 %
Kappa statistic                   0.8255
Mean absolute error               0.0213
Root mean squared error           0.1309
Relative absolute error           18.5583 %
Root relative squared error      55.5524 %
Total Number of Instances         1509

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.993     0.202     0.987     0.993     0.990     0.827     0.880     0.979     negative
      0.798     0.007     0.877     0.798     0.835     0.827     0.880     0.848     sick
Weighted Avg.      0.981     0.191     0.981     0.981     0.981     0.827     0.880     0.971

==== Confusion Matrix ====

      a      b  <-- classified as
1410   10 |      a = negative
      18   71 |      b = sick

```

## Percentage Split 66%

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B40-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  258
              [list of attributes omitted]
Test mode:   split 66.0% train, remainder test

==== Classifier model (full training set) ====

PART decision list
-----
T3_4 = '(1.105-inf)' AND
FTI_17 = '(-inf-169.025]' AND
TT4_5 = '(55.5-inf)': negative (3174.48/2.9)

referral_source = other AND
sick = f AND
age_6 = '(-inf-69.1)' AND
I131_treatment = f AND
TSH_measured = t AND
T4U_15 = '(1.02625-inf)': negative (69.64)

T3_measured = f: negative (106.58/2.1)

T3_10 = '(2.6875-inf)': negative (55.39)

TT4_5 = '(-inf-55.5]' AND
TSH_2 = '(26.50475-inf)': negative (37.13)

TT4_18 = '(-inf-194.6]' AND
referral_source = SVI AND
TSH_1 = '(-inf-13.254875)' AND
FTI_16 = '(-inf-159.2)' AND
TT4_5 = '(55.5-inf)' AND
TT4_6 = '(60.95-inf)': sick (156.47/0.58)

sick = f AND
TT4_18 = '(-inf-194.6]' AND
age_1 = '(12.35-inf)' AND
T4U_measured = t AND
T4U_4 = '(0.457-inf)' AND
T4U_5 = '(0.50875-inf)' AND
FTI_1 = '(11.825-inf)' AND
T4U_17 = '(-inf-1.12975]' AND
FTI_4 = '(41.3-inf)' AND
TSH_2 = '(-inf-26.50475]' AND
T3_9 = '(-inf-2.42375)' AND
referral_source = other AND
T3_2 = '(0.5775-inf)' AND
query_hypothyroid = f AND
T4U_11 = '(-inf-0.81925)': negative (16.9)

sick = f AND
TT4_18 = '(-inf-194.6]' AND
age_1 = '(12.35-inf)' AND
T4U_measured = f: negative (12.02/0.93)

sick = f AND
TT4_18 = '(-inf-194.6]' AND
T4U_4 = '(0.457-inf)' AND
TT4_1 = '(12.7-inf)' AND
age_1 = '(12.35-inf)' AND
T4U_17 = '(-inf-1.12975)' AND
FTI_4 = '(41.3-inf)' AND
T4U_5 = '(0.50875-inf)' AND
TSH_2 = '(-inf-26.50475)' AND
T3_9 = '(-inf-2.42375)' AND
age_3 = '(-inf-35.05)' AND
T3_8 = '(-inf-2.16)': sick (9.0/1.0)

sick = f AND
TT4_18 = '(-inf-194.6)' AND
age_2 = '(23.7-inf)' AND
FTI_1 = '(11.825-inf)' AND
T4U_3 = '(0.40525-inf)' AND
T4U_5 = '(0.50875-inf)' AND
T4U_17 = '(-inf-1.12975)' AND
FTI_4 = '(41.3-inf)' AND
TSH_2 = '(-inf-26.50475)' AND
age_3 = '(35.05-inf)' AND
T3_8 = '(-inf-2.16)' AND
FTI_15 = '(149.375-inf)' AND

```

```

TT4_16 = '(-inf-173.2]': negative (24.04/3.0)
TT4_18 = '(194.6-inf)': negative (8.0)

age_2 = '(23.7-inf)' AND
TT4_1 = '(12.7-inf)' AND
T4U_3 = '(0.40525-inf)' AND
T4U_17 = '(-inf-1.12975]' AND
age_7 = '(-inf-80.45]' AND
sick = f AND
FTI_20 = '(-inf-198.5]' AND
FTI_4 = '(41.3-inf)' AND
T4U_5 = '(0.50875-inf)' AND
TSH_2 = '(-inf-26.50475]' AND
age_3 = '(35.05-inf)' AND
T3_8 = '(-inf-2.16]' AND
FTI_19 = '(-inf-188.675]' AND
FTI_13 = '(-inf-129.725]' AND
T3_6 = '(-inf-1.6325]' AND
psych = f AND
referral_source = other AND
age_5 = '(-inf-57.75]': negative (7.03)

age_2 = '(23.7-inf)' AND
TT4_1 = '(12.7-inf)' AND
T4U_3 = '(0.40525-inf)' AND
T4U_17 = '(-inf-1.12975]' AND
TT4_6 = '(66.2-inf)' AND
age_3 = '(35.05-inf)': sick (42.24/4.07)

FTI_14 = '(-inf-139.55]' AND
age_7 = '(-inf-80.45]' AND
FTI_12 = '(-inf-119.9]' AND
TT4_1 = '(12.7-inf)' AND
query_hypothyroid = f AND
on_thyroxine = f AND
FTI_8 = '(-inf-80.6]' AND
query_hyperthyroid = f AND
FTI_4 = '(41.3-inf)' AND
T3_6 = '(-inf-1.6325]' AND
age_4 = '(46.4-inf)' AND
sex = F AND
T4U_6 = '(0.5605-inf)': negative (9.0/1.0)

FTI_14 = '(-inf-139.55]' AND
query_hypothyroid = f AND
FTI_1 = '(11.825-inf)' AND
age_2 = '(23.7-inf)' AND
query_hyperthyroid = f AND
T3_6 = '(-inf-1.6325]' AND
T4U_8 = '(0.664-inf)': sick (15.0/1.0)

age_4 = '(-inf-46.4]': negative (10.05/0.01)

TT4_5 = '(-inf-55.5]' AND
TT4_4 = '(44.8-inf)': negative (8.01)

FTI_12 = '(-inf-119.9]' AND
referral_source = SVI AND
age_5 = '(57.75-inf)': negative (4.0/1.0)

TSH_1 = '(-inf-13.254875)': sick (5.01/1.01)

: negative (2.0)

Number of Rules : 20

Time taken to build model: 0.72 seconds
==== Evaluation on test split ====
Time taken to test model on test split: 0.01 seconds
==== Summary ====
Correctly Classified Instances      1261          98.3619 %
Incorrectly Classified Instances    21           1.6381 %
Kappa statistic                   0.8445
Mean absolute error               0.0191
Root mean squared error           0.1235
Relative absolute error           16.7649 %
Root relative squared error      52.6281 %
Total Number of Instances         1282

==== Detailed Accuracy By Class ====

```

```

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
      0.994    0.187    0.988     0.994    0.991     0.846    0.919     0.987    negative
      0.813    0.006    0.897     0.813    0.853     0.846    0.919     0.811    sick
Weighted Avg.    0.984    0.176    0.983     0.984    0.983     0.846    0.919     0.976

==== Confusion Matrix ====
a      b  <-- classified as
1200   7 |   a = negative
  14  61 |   b = sick

```

## Leave one out fold

```

==== Run information ====
Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B40-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  258
              [list of attributes omitted]
Test mode:   3772-fold cross-validation

==== Classifier model (full training set) ====
PART decision list
-----
T3_4 = '(1.105-inf)' AND
FTI_17 = '(-inf-169.025]' AND
TT4_5 = '(55.5-inf)': negative (3174.48/2.9)

referral_source = other AND
sick = f AND
age_6 = '(-inf-69.1]' AND
I131_treatment = f AND
TSH_measured = t AND
T4U_15 = '(1.02625-inf)': negative (69.64)

T3_measured = f: negative (106.58/2.1)

T3_10 = '(2.6875-inf)': negative (55.39)

TT4_5 = '(-inf-55.5]' AND
TSH_2 = '(26.50475-inf)': negative (37.13)

TT4_18 = '(-inf-194.6]' AND
referral_source = SVI AND
TSH_1 = '(-inf-13.254875)' AND
FTI_16 = '(-inf-159.2)' AND
TT4_5 = '(55.5-inf)' AND
FTI_6 = '(60.95-inf)': sick (156.47/0.58)

sick = f AND
TT4_18 = '(-inf-194.6)' AND
age_1 = '(12.35-inf)' AND
T4U_measured = t AND
T4U_4 = '(0.457-inf)' AND
T4U_5 = '(0.50875-inf)' AND
FTI_1 = '(11.825-inf)' AND
T4U_17 = '(-inf-1.12975]' AND
FTI_4 = '(41.3-inf)' AND
TSH_2 = '(-inf-26.50475]' AND
T3_9 = '(-inf-2.42375)' AND
referral_source = other AND
T3_2 = '(0.5775-inf)' AND
query_hypothyroid = f AND
T4U_11 = '(-inf-0.81925)': negative (16.9)

sick = f AND
TT4_18 = '(-inf-194.6)' AND
age_1 = '(12.35-inf)' AND
T4U_measured = f: negative (12.02/0.93)

sick = f AND
TT4_18 = '(-inf-194.6)' AND
T4U_4 = '(0.457-inf)' AND
TT4_1 = '(12.7-inf)' AND
age_1 = '(12.35-inf)' AND
T4U_17 = '(-inf-1.12975)' AND
FTI_4 = '(41.3-inf)' AND
T4U_5 = '(0.50875-inf)' AND
TSH_2 = '(-inf-26.50475)' AND
T3_9 = '(-inf-2.42375)' AND
age_3 = '(-inf-35.05)' AND

```

```

T3_8 = '(-inf-2.16)': sick (9.0/1.0)

sick = f AND
TT4_18 = '(-inf-194.6)' AND
age_2 = '(23.7-inf)' AND
FTI_1 = '(11.825-inf)' AND
T4U_3 = '(0.40525-inf)' AND
T4U_5 = '(0.50875-inf)' AND
T4U_17 = '(-inf-1.12975)' AND
FTI_4 = '(41.3-inf)' AND
TSH_2 = '(-inf-26.50475)' AND
age_3 = '(35.05-inf)' AND
T3_8 = '(-inf-2.16)' AND
FTI_15 = '(149.375-inf)' AND
TT4_16 = '(-inf-173.2)': negative (24.04/3.0)

TT4_18 = '(194.6-inf)': negative (8.0)

age_2 = '(23.7-inf)' AND
TT4_1 = '(12.7-inf)' AND
T4U_3 = '(0.40525-inf)' AND
T4U_17 = '(-inf-1.12975)' AND
age_7 = '(-inf-80.45)' AND
sick = f AND
FTI_20 = '(-inf-198.5)' AND
FTI_4 = '(41.3-inf)' AND
T4U_5 = '(0.50875-inf)' AND
TSH_2 = '(-inf-26.50475)' AND
age_3 = '(35.05-inf)' AND
T3_8 = '(-inf-2.16)' AND
FTI_19 = '(-inf-188.675)' AND
FTI_13 = '(-inf-129.725)' AND
T3_6 = '(-inf-1.6325)' AND
psych = f AND
referral_source = other AND
age_5 = '(-inf-57.75)': negative (7.03)

age_2 = '(23.7-inf)' AND
TT4_1 = '(12.7-inf)' AND
T4U_3 = '(0.40525-inf)' AND
T4U_17 = '(-inf-1.12975)' AND
TT4_6 = '(66.2-inf)' AND
age_3 = '(35.05-inf)': sick (42.24/4.07)

FTI_14 = '(-inf-139.55)' AND
age_7 = '(-inf-80.45)' AND
FTI_12 = '(-inf-119.9)' AND
TT4_1 = '(12.7-inf)' AND
query_hypothyroid = f AND
on_thyroxine = f AND
FTI_8 = '(-inf-80.6)' AND
query_hyperthyroid = f AND
FTI_4 = '(41.3-inf)' AND
T3_6 = '(-inf-1.6325)' AND
age_4 = '(46.4-inf)' AND
sex = F AND
T4U_6 = '(0.5605-inf)': negative (9.0/1.0)

FTI_14 = '(-inf-139.55)' AND
query_hypothyroid = f AND
FTI_1 = '(11.825-inf)' AND
age_2 = '(23.7-inf)' AND
query_hyperthyroid = f AND
T3_6 = '(-inf-1.6325)' AND
T4U_8 = '(0.664-inf)': sick (15.0/1.0)

age_4 = '(-inf-46.4)': negative (10.05/0.01)

TT4_5 = '(-inf-55.5)' AND
TT4_4 = '(44.8-inf)': negative (8.01)

FTI_12 = '(-inf-119.9)' AND
referral_source = SVI AND
age_5 = '(57.75-inf)': negative (4.0/1.0)

TSH_1 = '(-inf-13.254875)': sick (5.01/1.01)

: negative (2.0)

Number of Rules : 20

```

Time taken to build model: 0.76 seconds

```

==== Stratified cross-validation ====
==== Summary ====

```

```

Correctly Classified Instances      3720          98.6214 %
Incorrectly Classified Instances   52           1.3786 %
Kappa statistic                   0.8791
Mean absolute error               0.0176
Root mean squared error          0.1142
Relative absolute error           15.2342 %
Root relative squared error      47.6098 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall    F-Measure   MCC     ROC Area   PRC Area   Class
           0.993     0.121     0.992     0.993     0.993     0.879     0.929     0.987   negative
           0.879     0.007     0.894     0.879     0.886     0.879     0.929     0.861   sick
Weighted Avg.       0.986     0.114     0.986     0.986     0.986     0.879     0.929     0.980

==== Confusion Matrix ====

      a      b  <-- classified as
3517  24 |      a = negative
  28  203 |      b = sick

```

### 3.1.5 Conclusion

PART							
Preprocessing method	Cross Validation			Percentage Split			Leave One Out Fold
	5 folds	10 folds	15 folds	50.00%	60.00%	66.00%	
With Original Dataset	98.3033%	98.6214%	98.5949%	98.6214%	97.8131%	98.7520%	98.8865%
Equal width binning	40 bins	98.3298%	98.4889%	98.4093%	97.7731%	98.1445%	98.3619%
	10 bins	98.2238%	97.9321%	98.1177%	97.4549%	98.0782%	97.8939%
	5 bins	93.8494%	94.1676%	94.1145%	93.7964%	94.0358%	94.4618%
	2 bins	93.7434%	93.8494%	93.8229%	93.7434%	94.1021%	93.2917%

PART was able to benefit from **discretizing with equal width binning and 40 bins** in two tests in Cross validation 5 folds and in Percentage split 60%. However the benefits are very negligible.

## 3.2 Non supervised discretization and Equal frequency binning

### 3.2.1 2 bins

#### Cross Validation 5 folds

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B2-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  30
             age_1
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             I131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary
             psych
             TSH_measured
             TSH_1
             T3_measured
             T3_1
             TT4_measured

```

```

TT4_1
T4U_measured
T4U_1
FTI_measured
FTI_1
TBG_measured
TBG
referral_source
Class
Test mode: 5-fold cross-validation

==== Classifier model (full training set) ====

PART decision list
-----
referral_source = other: negative (2201.0/33.0)

T3_1 = '(1.95-inf)': negative (699.39/6.89)

T3_measured = t AND
referral_source = SVHC AND
query_hypothyroid = f AND
FTI_1 = '(106.5-inf)': negative (65.42)

T3_measured = t AND
thyroid_surgery = f AND
T4U_1 = '(0.975-inf)' AND
query_hypothyroid = f: negative (204.2/18.05)

T3_measured = t AND
thyroid_surgery = f AND
TT4_1 = '(103.5-inf)': negative (160.32/32.83)

T3_measured = f: negative (41.61/1.11)

thyroid_surgery = f AND
query_hypothyroid = f AND
referral_source = SVHC AND
age_1 = '(-inf-54.5]': negative (24.74/1.0)

thyroid_surgery = f AND
query_hypothyroid = f AND
referral_source = SVHD: negative (8.0/1.0)

thyroid_surgery = f AND
query_hypothyroid = f AND
T4U_measured = f: negative (10.61/2.12)

thyroid_surgery = f AND
T4U_1 = '(-inf-0.975]' AND
sick = f AND
referral_source = SVHC AND
on_thyroxine = f AND
sex = M: negative (10.63/1.0)

thyroid_surgery = t: negative (8.0)

T4U_1 = '(0.975-inf)': sick (5.01/1.0)

psych = t AND
sex = F: negative (5.0/2.0)

psych = f AND
referral_source = SVI AND
sick = f AND
on_thyroxine = f AND
query_hyperthyroid = f AND
query_hypothyroid = f AND
age_1 = '(54.5-inf)' AND
FTI_1 = '(-inf-106.5)': negative (113.0/38.0)

psych = f AND
referral_source = SVI AND
age_1 = '(-inf-54.5]': negative (78.0/25.0)

psych = f AND
age_1 = '(54.5-inf)' AND
referral_source = SVI AND
query_hyperthyroid = f AND
sex = M AND
query_hypothyroid = f: negative (43.23/15.41)

psych = f AND
age_1 = '(54.5-inf)' AND
referral_source = SVI AND
query_hyperthyroid = f AND
query_hypothyroid = f

```

```

FTI_1 = '(-inf-106.5]' AND
TSH_1 = '(1.35-inf)' AND
sick = f: negative (11.0/3.0)

psych = f AND
age_1 = '(54.5-inf)' AND
sex = M: sick (6.9/0.85)

query_hyperthyroid = f AND
psych = f AND
referral_source = SVHC: sick (4.34/1.34)

query_hyperthyroid = f AND
psych = f AND
FTI_1 = '(-inf-106.5]' AND
TSH_1 = '(1.35-inf)': sick (7.0/3.0)

query_hyperthyroid = f AND
psych = f AND
FTI_1 = '(106.5-inf)': sick (54.61/25.07)

sex = F: negative (6.0)

: sick (3.99)

Number of Rules : 23

Time taken to build model: 0.12 seconds

==== Stratified cross-validation ====
==== Summary ====

    Correctly Classified Instances      3537           93.7699 %
    Incorrectly Classified Instances   235            6.2301 %
    Kappa statistic                   0.1915
    Mean absolute error              0.0887
    Root mean squared error          0.2157
    Relative absolute error          76.9899 %
    Root relative squared error     89.9448 %
    Total Number of Instances       3772

==== Detailed Accuracy By Class ====

    TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
    0.990    0.861    0.946      0.990    0.968      0.231    0.859      0.983    negative
    0.139    0.010    0.471      0.139    0.214      0.231    0.859      0.352    sick
    Weighted Avg.      0.938    0.809    0.917      0.938    0.921      0.231    0.859      0.944

==== Confusion Matrix ====

    a      b  <-- classified as
3505   36 |    a = negative
  199   32 |    b = sick

```

## Cross Validation 10 folds

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B2-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  30
             age_1
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             I131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary
             psych
             TSH_measured
             TSH_1
             T3_measured
             T3_1
             TT4_measured

```

```

TT4_1
T4U_measured
T4U_1
FTI_measured
FTI_1
TBG_measured
TBG
referral_source
Class
Test mode: 10-fold cross-validation

==== Classifier model (full training set) ====

PART decision list
-----
referral_source = other: negative (2201.0/33.0)

T3_1 = '(1.95-inf)': negative (699.39/6.89)

T3_measured = t AND
referral_source = SVHC AND
query_hypothyroid = f AND
FTI_1 = '(106.5-inf)': negative (65.42)

T3_measured = t AND
thyroid_surgery = f AND
T4U_1 = '(0.975-inf)' AND
query_hypothyroid = f: negative (204.2/18.05)

T3_measured = t AND
thyroid_surgery = f AND
TT4_1 = '(103.5-inf)': negative (160.32/32.83)

T3_measured = f: negative (41.61/1.11)

thyroid_surgery = f AND
query_hypothyroid = f AND
referral_source = SVHC AND
age_1 = '(-inf-54.5]': negative (24.74/1.0)

thyroid_surgery = f AND
query_hypothyroid = f AND
referral_source = SVHD: negative (8.0/1.0)

thyroid_surgery = f AND
query_hypothyroid = f AND
T4U_measured = f: negative (10.61/2.12)

thyroid_surgery = f AND
T4U_1 = '(-inf-0.975]' AND
sick = f AND
referral_source = SVHC AND
on_thyroxine = f AND
sex = M: negative (10.63/1.0)

thyroid_surgery = t: negative (8.0)

T4U_1 = '(0.975-inf)': sick (5.01/1.0)

psych = t AND
sex = F: negative (5.0/2.0)

psych = f AND
referral_source = SVI AND
sick = f AND
on_thyroxine = f AND
query_hyperthyroid = f AND
query_hypothyroid = f AND
age_1 = '(54.5-inf)' AND
FTI_1 = '(-inf-106.5)': negative (113.0/38.0)

psych = f AND
referral_source = SVI AND
age_1 = '(-inf-54.5]': negative (78.0/25.0)

psych = f AND
age_1 = '(54.5-inf)' AND
referral_source = SVI AND
query_hyperthyroid = f AND
sex = M AND
query_hypothyroid = f: negative (43.23/15.41)

psych = f AND
age_1 = '(54.5-inf)' AND
referral_source = SVI AND
query_hyperthyroid = f AND
query_hypothyroid = f

```

```

FTI_1 = '(-inf-106.5]' AND
TSH_1 = '(1.35-inf)' AND
sick = f: negative (11.0/3.0)

psych = f AND
age_1 = '(54.5-inf)' AND
sex = M: sick (6.9/0.85)

query_hyperthyroid = f AND
psych = f AND
referral_source = SVHC: sick (4.34/1.34)

query_hyperthyroid = f AND
psych = f AND
FTI_1 = '(-inf-106.5]' AND
TSH_1 = '(1.35-inf)': sick (7.0/3.0)

query_hyperthyroid = f AND
psych = f AND
FTI_1 = '(106.5-inf)': sick (54.61/25.07)

sex = F: negative (6.0)

: sick (3.99)

Number of Rules : 23

Time taken to build model: 0.12 seconds

==== Stratified cross-validation ====
==== Summary ====

    Correctly Classified Instances      3525           93.4517 %
    Incorrectly Classified Instances   247            6.5483 %
    Kappa statistic                   0.1238
    Mean absolute error              0.0903
    Root mean squared error          0.2187
    Relative absolute error          78.3432 %
    Root relative squared error     91.2039 %
    Total Number of Instances       3772

==== Detailed Accuracy By Class ====

    TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
    0.990     0.909     0.943      0.990     0.966      0.157     0.846      0.980     negative
    0.091     0.010     0.362      0.091     0.145      0.157     0.846      0.321     sick
    Weighted Avg.   0.935     0.854     0.908      0.935     0.916      0.157     0.846      0.940

==== Confusion Matrix ====

    a      b  <-- classified as
3504    37 |    a = negative
  210    21 |    b = sick

```

## Cross Validation 15 folds

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B2-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  30
             age_1
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             I131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary
             psych
             TSH_measured
             TSH_1
             T3_measured
             T3_1
             TT4_measured

```

```

TT4_1
T4U_measured
T4U_1
FTI_measured
FTI_1
TBG_measured
TBG
referral_source
Class
Test mode: 15-fold cross-validation

==== Classifier model (full training set) ====

PART decision list
-----
referral_source = other: negative (2201.0/33.0)

T3_1 = '(1.95-inf)': negative (699.39/6.89)

T3_measured = t AND
referral_source = SVHC AND
query_hypothyroid = f AND
FTI_1 = '(106.5-inf)': negative (65.42)

T3_measured = t AND
thyroid_surgery = f AND
T4U_1 = '(0.975-inf)' AND
query_hypothyroid = f: negative (204.2/18.05)

T3_measured = t AND
thyroid_surgery = f AND
TT4_1 = '(103.5-inf)': negative (160.32/32.83)

T3_measured = f: negative (41.61/1.11)

thyroid_surgery = f AND
query_hypothyroid = f AND
referral_source = SVHC AND
age_1 = '(-inf-54.5]': negative (24.74/1.0)

thyroid_surgery = f AND
query_hypothyroid = f AND
referral_source = SVHD: negative (8.0/1.0)

thyroid_surgery = f AND
query_hypothyroid = f AND
T4U_measured = f: negative (10.61/2.12)

thyroid_surgery = f AND
T4U_1 = '(-inf-0.975]' AND
sick = f AND
referral_source = SVHC AND
on_thyroxine = f AND
sex = M: negative (10.63/1.0)

thyroid_surgery = t: negative (8.0)

T4U_1 = '(0.975-inf)': sick (5.01/1.0)

psych = t AND
sex = F: negative (5.0/2.0)

psych = f AND
referral_source = SVI AND
sick = f AND
on_thyroxine = f AND
query_hyperthyroid = f AND
query_hypothyroid = f AND
age_1 = '(54.5-inf)' AND
FTI_1 = '(-inf-106.5)': negative (113.0/38.0)

psych = f AND
referral_source = SVI AND
age_1 = '(-inf-54.5]': negative (78.0/25.0)

psych = f AND
age_1 = '(54.5-inf)' AND
referral_source = SVI AND
query_hyperthyroid = f AND
sex = M AND
query_hypothyroid = f: negative (43.23/15.41)

psych = f AND
age_1 = '(54.5-inf)' AND
referral_source = SVI AND
query_hyperthyroid = f AND
query_hypothyroid = f

```

```

FTI_1 = '(-inf-106.5]' AND
TSH_1 = '(1.35-inf)' AND
sick = f: negative (11.0/3.0)

psych = f AND
age_1 = '(54.5-inf)' AND
sex = M: sick (6.9/0.85)

query_hyperthyroid = f AND
psych = f AND
referral_source = SVHC: sick (4.34/1.34)

query_hyperthyroid = f AND
psych = f AND
FTI_1 = '(-inf-106.5]' AND
TSH_1 = '(1.35-inf)': sick (7.0/3.0)

query_hyperthyroid = f AND
psych = f AND
FTI_1 = '(106.5-inf)': sick (54.61/25.07)

sex = F: negative (6.0)

: sick (3.99)

Number of Rules : 23

Time taken to build model: 0.08 seconds

==== Stratified cross-validation ====
==== Summary ====

    Correctly Classified Instances      3536           93.7434 %
    Incorrectly Classified Instances   236            6.2566 %
    Kappa statistic                   0.2002
    Mean absolute error              0.0895
    Root mean squared error          0.2168
    Relative absolute error          77.6626 %
    Root relative squared error     90.4145 %
    Total Number of Instances       3772

==== Detailed Accuracy By Class ====

    TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
    0.989     0.853     0.947      0.989     0.967      0.237    0.847      0.982     negative
    0.147     0.011     0.466      0.147     0.224      0.237    0.847      0.337     sick
    Weighted Avg.   0.937     0.801     0.917      0.937     0.922      0.237    0.847      0.943

==== Confusion Matrix ====

    a      b  <-- classified as
3502   39 |   a = negative
  197   34 |   b = sick

```

## Percentage Split 50%

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B2-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  30
             age_1
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             I131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary
             psych
             TSH_measured
             TSH_1
             T3_measured
             T3_1
             TT4_measured

```

```

TT4_1
T4U_measured
T4U_1
FTI_measured
FTI_1
TBG_measured
TBG
referral_source
Class
Test mode: split 50.0% train, remainder test

==== Classifier model (full training set) ===

PART decision list
-----
referral_source = other: negative (2201.0/33.0)

T3_1 = '(1.95-inf)': negative (699.39/6.89)

T3_measured = t AND
referral_source = SVHC AND
query_hypothyroid = f AND
FTI_1 = '(106.5-inf)': negative (65.42)

T3_measured = t AND
thyroid_surgery = f AND
T4U_1 = '(0.975-inf)' AND
query_hypothyroid = f: negative (204.2/18.05)

T3_measured = t AND
thyroid_surgery = f AND
TT4_1 = '(103.5-inf)': negative (160.32/32.83)

T3_measured = f: negative (41.61/1.11)

thyroid_surgery = f AND
query_hypothyroid = f AND
referral_source = SVHC AND
age_1 = '(-inf-54.5]': negative (24.74/1.0)

thyroid_surgery = f AND
query_hypothyroid = f AND
referral_source = SVHD: negative (8.0/1.0)

thyroid_surgery = f AND
query_hypothyroid = f AND
T4U_measured = f: negative (10.61/2.12)

thyroid_surgery = f AND
T4U_1 = '(-inf-0.975]' AND
sick = f AND
referral_source = SVHC AND
on_thyroxine = f AND
sex = M: negative (10.63/1.0)

thyroid_surgery = t: negative (8.0)

T4U_1 = '(0.975-inf)': sick (5.01/1.0)

psych = t AND
sex = F: negative (5.0/2.0)

psych = f AND
referral_source = SVI AND
sick = f AND
on_thyroxine = f AND
query_hyperthyroid = f AND
query_hypothyroid = f AND
age_1 = '(54.5-inf)' AND
FTI_1 = '(-inf-106.5)': negative (113.0/38.0)

psych = f AND
referral_source = SVI AND
age_1 = '(-inf-54.5]': negative (78.0/25.0)

psych = f AND
age_1 = '(54.5-inf)' AND
referral_source = SVI AND
query_hyperthyroid = f AND
sex = M AND
query_hypothyroid = f: negative (43.23/15.41)

psych = f AND
age_1 = '(54.5-inf)' AND
referral_source = SVI AND
query_hyperthyroid = f AND
query_hypothyroid = f

```

```

FTI_1 = '(-inf-106.5]' AND
TSH_1 = '(1.35-inf)' AND
sick = f: negative (11.0/3.0)

psych = f AND
age_1 = '(54.5-inf)' AND
sex = M: sick (6.9/0.85)

query_hyperthyroid = f AND
psych = f AND
referral_source = SVHC: sick (4.34/1.34)

query_hyperthyroid = f AND
psych = f AND
FTI_1 = '(-inf-106.5]' AND
TSH_1 = '(1.35-inf)': sick (7.0/3.0)

query_hyperthyroid = f AND
psych = f AND
FTI_1 = '(106.5-inf)': sick (54.61/25.07)

sex = F: negative (6.0)

: sick (3.99)

Number of Rules : 23

Time taken to build model: 0.11 seconds

==== Evaluation on test split ====

Time taken to test model on test split: 0 seconds

==== Summary ====

    Correctly Classified Instances      1766           93.6373 %
    Incorrectly Classified Instances   120            6.3627 %
    Kappa statistic                   0.1676
    Mean absolute error              0.0874
    Root mean squared error          0.2228
    Relative absolute error          75.722 %
    Root relative squared error     92.3569 %
    Total Number of Instances        1886

==== Detailed Accuracy By Class ====

    TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
    0.990     0.880     0.944      0.990     0.967      0.209     0.783      0.970     negative
    0.120     0.010     0.452      0.120     0.189      0.209     0.783      0.284     sick
    Weighted Avg.   0.936     0.826     0.914      0.936     0.919      0.209     0.783      0.928

==== Confusion Matrix ====

    a      b  <-- classified as
1752   17 |      a = negative
  103   14 |      b = sick

```

## Percentage Split 60%

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B2-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  30
             age_1
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             I131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary
             psych
             TSH_measured
             TSH_1

```

```

T3_measured
T3_1
TT4_measured
TT4_1
T4U_measured
T4U_1
FTI_measured
FTI_1
TBG_measured
TBG
referral_source
Class
Test mode: split 60.0% train, remainder test

==== Classifier model (full training set) ====

PART decision list
-----
referral_source = other: negative (2201.0/33.0)

T3_1 = '(1.95-inf)': negative (699.39/6.89)

T3_measured = t AND
referral_source = SVHC AND
query_hypothyroid = f AND
FTI_1 = '(106.5-inf)': negative (65.42)

T3_measured = t AND
thyroid_surgery = f AND
T4U_1 = '(0.975-inf)' AND
query_hypothyroid = f: negative (204.2/18.05)

T3_measured = t AND
thyroid_surgery = f AND
TT4_1 = '(103.5-inf)': negative (160.32/32.83)

T3_measured = f: negative (41.61/1.11)

thyroid_surgery = f AND
query_hypothyroid = f AND
referral_source = SVHC AND
age_1 = '(-inf-54.5)': negative (24.74/1.0)

thyroid_surgery = f AND
query_hypothyroid = f AND
referral_source = SVHD: negative (8.0/1.0)

thyroid_surgery = f AND
query_hypothyroid = f AND
T4U_measured = f: negative (10.61/2.12)

thyroid_surgery = f AND
T4U_1 = '(-inf-0.975)' AND
sick = f AND
referral_source = SVHC AND
on_thyroxine = f AND
sex = M: negative (10.63/1.0)

thyroid_surgery = t: negative (8.0)

T4U_1 = '(0.975-inf)': sick (5.01/1.0)

psych = t AND
sex = F: negative (5.0/2.0)

psych = f AND
referral_source = SVI AND
sick = f AND
on_thyroxine = f AND
query_hyperthyroid = f AND
query_hypothyroid = f AND
age_1 = '(54.5-inf)' AND
FTI_1 = '(-inf-106.5)': negative (113.0/38.0)

psych = f AND
referral_source = SVI AND
age_1 = '(-inf-54.5)': negative (78.0/25.0)

psych = f AND
age_1 = '(54.5-inf)' AND
referral_source = SVI AND
query_hyperthyroid = f AND
sex = M AND
query_hypothyroid = f: negative (43.23/15.41)

psych = f AND

```

```

age_1 = '(54.5-inf)' AND
referral_source = SVI AND
query_hyperthyroid = f AND
FTI_1 = '(-inf-106.5]' AND
TSH_1 = '(1.35-inf)' AND
sick = f: negative (11.0/3.0)

psych = f AND
age_1 = '(54.5-inf)' AND
sex = M: sick (6.9/0.85)

query_hyperthyroid = f AND
psych = f AND
referral_source = SVHC: sick (4.34/1.34)

query_hyperthyroid = f AND
psych = f AND
FTI_1 = '(-inf-106.5]' AND
TSH_1 = '(1.35-inf)': sick (7.0/3.0)

query_hyperthyroid = f AND
psych = f AND
FTI_1 = '(106.5-inf)': sick (54.61/25.07)

sex = F: negative (6.0)

: sick (3.99)

Number of Rules : 23

Time taken to build model: 0.09 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

==== Summary ===

    Correctly Classified Instances      1417           93.9032 %
    Incorrectly Classified Instances   92            6.0968 %
    Kappa statistic                   0.0151
    Mean absolute error              0.0839
    Root mean squared error          0.214
    Relative absolute error          73.1489 %
    Root relative squared error     90.8143 %
    Total Number of Instances       1509

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.997     0.989     0.941      0.997     0.969     0.035     0.835      0.981   negative
      0.011     0.003     0.200      0.011     0.021     0.035     0.835      0.317   sick
Weighted Avg.   0.939     0.931     0.898      0.939     0.913     0.035     0.835      0.942

==== Confusion Matrix ===

      a     b  <-- classified as
1416    4 |     a = negative
      88   1 |     b = sick

```

## Percentage Split 66%

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B2-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  30
             age_1
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             i131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary

```

```

psych
TSH_measured
TSH_1
T3_measured
T3_1
TT4_measured
TT4_1
T4U_measured
T4U_1
FTI_measured
FTI_1
TBG_measured
TBG
referral_source
Class
Test mode: split 66.0% train, remainder test

==== Classifier model (full training set) ===

PART decision list
-----
referral_source = other: negative (2201.0/33.0)

T3_1 = '(1.95-inf)': negative (699.39/6.89)

T3_measured = t AND
referral_source = SVHC AND
query_hypothyroid = f AND
FTI_1 = '(106.5-inf)': negative (65.42)

T3_measured = t AND
thyroid_surgery = f AND
T4U_1 = '(0.975-inf)' AND
query_hypothyroid = f: negative (204.2/18.05)

T3_measured = t AND
thyroid_surgery = f AND
TT4_1 = '(103.5-inf)': negative (160.32/32.83)

T3_measured = f: negative (41.61/1.11)

thyroid_surgery = f AND
query_hypothyroid = f AND
referral_source = SVHC AND
age_1 = '(-inf-54.5]': negative (24.74/1.0)

thyroid_surgery = f AND
query_hypothyroid = f AND
referral_source = SVHD: negative (8.0/1.0)

thyroid_surgery = f AND
query_hypothyroid = f AND
T4U_measured = f: negative (10.61/2.12)

thyroid_surgery = f AND
T4U_1 = '(-inf-0.975]' AND
sick = f AND
referral_source = SVHC AND
on_thyroxine = f AND
sex = M: negative (10.63/1.0)

thyroid_surgery = t: negative (8.0)

T4U_1 = '(0.975-inf)': sick (5.01/1.0)

psych = t AND
sex = F: negative (5.0/2.0)

psych = f AND
referral_source = SVI AND
sick = f AND
on_thyroxine = f AND
query_hyperthyroid = f AND
query_hypothyroid = f AND
age_1 = '(54.5-inf)' AND
FTI_1 = '(-inf-106.5)': negative (113.0/38.0)

psych = f AND
referral_source = SVI AND
age_1 = '(-inf-54.5]': negative (78.0/25.0)

psych = f AND
age_1 = '(54.5-inf)' AND
referral_source = SVI AND
query_hyperthyroid = f AND
sex = M AND

```

```

query_hypothyroid = f: negative (43.23/15.41)

psych = f AND
age_1 = '(54.5-inf)' AND
referral_source = SVI AND
query_hyperthyroid = f AND
FTI_1 = '(-inf-106.5]' AND
TSH_1 = '(1.35-inf)' AND
sick = f: negative (11.0/3.0)

psych = f AND
age_1 = '(54.5-inf)' AND
sex = M: sick (6.9/0.85)

query_hyperthyroid = f AND
psych = f AND
referral_source = SVHC: sick (4.34/1.34)

query_hyperthyroid = f AND
psych = f AND
FTI_1 = '(-inf-106.5]' AND
TSH_1 = '(1.35-inf)': sick (7.0/3.0)

query_hyperthyroid = f AND
psych = f AND
FTI_1 = '(106.5-inf)': sick (54.61/25.07)

sex = F: negative (6.0)

: sick (3.99)

Number of Rules : 23

Time taken to build model: 0.09 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

==== Summary ===

    Correctly Classified Instances      1206           94.0718 %
    Incorrectly Classified Instances   76            5.9282 %
    Kappa statistic                   0.1875
    Mean absolute error              0.0845
    Root mean squared error          0.2153
    Relative absolute error          74.0655 %
    Root relative squared error     91.7292 %
    Total Number of Instances        1282

==== Detailed Accuracy By Class ===

    TP Rate   FP Rate   Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
    0.991     0.867    0.948     0.991    0.969     0.230    0.840     0.982    negative
    0.133     0.009    0.476     0.133    0.208     0.230    0.840     0.308    sick
    Weighted Avg. 0.941     0.816    0.921     0.941    0.925     0.230    0.840     0.943

==== Confusion Matrix ===

    a      b  <-- classified as
1196   11 |   a = negative
    65   10 |   b = sick

```

## Leave one out fold

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B2-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  30
             age_1
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             I131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium

```

```

goitre
tumor
hypopituitary
psych
TSH_measured
TSH_1
T3_measured
T3_1
TT4_measured
TT4_1
T4U_measured
T4U_1
FTI_measured
FTI_1
TBG_measured
TBG
referral_source
Class
Test mode: 3772-fold cross-validation

==== Classifier model (full training set) ====
PART decision list
-----
referral_source = other: negative (2201.0/33.0)

T3_1 = '(1.95-inf)': negative (699.39/6.89)

T3_measured = t AND
referral_source = SVHC AND
query_hypothyroid = f AND
FTI_1 = '(106.5-inf)': negative (65.42)

T3_measured = t AND
thyroid_surgery = f AND
T4U_1 = '(0.975-inf)': AND
query_hypothyroid = f: negative (204.2/18.05)

T3_measured = t AND
thyroid_surgery = f AND
TT4_1 = '(103.5-inf)': negative (160.32/32.83)

T3_measured = f: negative (41.61/1.11)

thyroid_surgery = f AND
query_hypothyroid = f AND
referral_source = SVHC AND
age_1 = '(-inf-54.5)': negative (24.74/1.0)

thyroid_surgery = f AND
query_hypothyroid = f AND
referral_source = SVHD: negative (8.0/1.0)

thyroid_surgery = f AND
query_hypothyroid = f AND
T4U_measured = f: negative (10.61/2.12)

thyroid_surgery = f AND
T4U_1 = '(-inf-0.975]': AND
sick = f AND
referral_source = SVHC AND
on_thyroxine = f AND
sex = M: negative (10.63/1.0)

thyroid_surgery = t: negative (8.0)

T4U_1 = '(0.975-inf)': sick (5.01/1.0)

psych = t AND
sex = F: negative (5.0/2.0)

psych = f AND
referral_source = SVI AND
sick = f AND
on_thyroxine = f AND
query_hyperthyroid = f AND
query_hypothyroid = f AND
age_1 = '(54.5-inf)': AND
FTI_1 = '(-inf-106.5)': negative (113.0/38.0)

psych = f AND
referral_source = SVI AND
age_1 = '(-inf-54.5)': negative (78.0/25.0)

psych = f AND
age_1 = '(54.5-inf)' AND

```

```

referral_source = SVI AND
query_hyperthyroid = f AND
sex = M AND
query_hypothyroid = f: negative (43.23/15.41)

psych = f AND
age_1 = '(54.5-inf)' AND
referral_source = SVI AND
query_hyperthyroid = f AND
FTI_1 = '(-inf-106.5]' AND
TSH_1 = '(1.35-inf)' AND
sick = f: negative (11.0/3.0)

psych = f AND
age_1 = '(54.5-inf)' AND
sex = M: sick (6.9/0.85)

query_hyperthyroid = f AND
psych = f AND
referral_source = SVHC: sick (4.34/1.34)

query_hyperthyroid = f AND
psych = f AND
FTI_1 = '(-inf-106.5]' AND
TSH_1 = '(1.35-inf)': sick (7.0/3.0)

query_hyperthyroid = f AND
psych = f AND
FTI_1 = '(106.5-inf)': sick (54.61/25.07)

sex = F: negative (6.0)

: sick (3.99)

Number of Rules : 23

Time taken to build model: 0.09 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3531          93.6108 %
Incorrectly Classified Instances   241           6.3892 %
Kappa statistic                   0.1709
Mean absolute error               0.0901
Root mean squared error          0.2189
Relative absolute error          78.1874 %
Root relative squared error     91.2584 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.989     0.874     0.945      0.989     0.967      0.206     0.787     0.979     negative
      0.126     0.011     0.426      0.126     0.194      0.206     0.787     0.326     sick
Weighted Avg.      0.936     0.822     0.914      0.936     0.919      0.206     0.787     0.939

==== Confusion Matrix ====

      a      b  <-- classified as
3502  39 |    a = negative
 202  29 |    b = sick

```

### 3.2.2 5 bins

#### Cross Validation 5 folds

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B5-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  48
              age_1
              age_2
              age_3
              age_4
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick

```

```

pregnant
thyroid_surgery
I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH_1
TSH_2
TSH_3
TSH_4
T3_measured
T3_1
T3_2
T3_3
T3_4
TT4_measured
TT4_1
TT4_2
TT4_3
TT4_4
T4U_measured
T4U_1
T4U_2
T4U_3
T4U_4
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
TBG_measured
TBG
referral_source
Class
Test mode: 5-fold cross-validation

==== Classifier model (full training set) ====

PART decision list
-----
T3_1 = '(1.47-inf)' AND
FTI_4 = '(-inf-129.5)': negative (2417.11/4.41)

referral_source = other AND
T3_measured = f: negative (265.4/0.2)

T3_1 = '(1.47-inf)' AND
TSH_1 = '(0.295-inf)': negative (230.71/0.17)

referral_source = other AND
sick = f AND
age_4 = '(-inf-70.5]' AND
T4U_4 = '(-inf-1.115]' AND
TSH_measured = t AND
T4U_measured = t AND
FTI_2 = '(101.5-inf)' AND
TSH_4 = '(-inf-3.15]' AND
on_thyroxine = t: negative (54.0)

referral_source = other AND
T4U_3 = '(1.015-inf)': negative (94.38/1.0)

T3_2 = '(1.85-inf)' AND
TT4_4 = '(131.5-inf)' AND
on_thyroxine = f AND
referral_source = other AND
T3_3 = '(2.15-inf)': negative (43.09/2.0)

T3_measured = t AND
T3_4 = '(2.45-inf)' AND
T4U_1 = '(0.855-inf)': negative (24.96)

T3_measured = t AND
thyroid_surgery = f AND
T3_2 = '(1.85-inf)' AND
TT4_4 = '(-inf-131.5)': negative (25.09)

T3_measured = t AND
T3_4 = '(-inf-2.45)' AND
thyroid_surgery = f AND
referral_source = other AND

```

```

T4U_measured = t AND
age_4 = '(-inf-70.5]' AND
TT4_4 = '(-inf-131.5]' AND
on_thyroxine = f AND
T4U_1 = '(0.855-inf)': negative (39.0)

T3_measured = f: negative (17.78/0.22)

T3_4 = '(-inf-2.45]' AND
thyroid_surgery = f AND
T3_1 = '(1.47-inf)' AND
on_thyroxine = f AND
TT4_4 = '(131.5-inf)' AND
age_1 = '(32.5-inf)' AND
sex = F AND
T4U_2 = '(0.935-inf)': negative (12.57/2.73)

T3_4 = '(-inf-2.45]' AND
thyroid_surgery = f AND
T3_1 = '(1.47-inf)' AND
on_thyroxine = t: negative (11.53)

T3_4 = '(-inf-2.45]' AND
thyroid_surgery = f AND
TT4_measured = t AND
sick = t AND
query_hypothyroid = f AND
T4U_3 = '(-inf-1.015]' AND
T4U_1 = '(-inf-0.855]' AND
FTI_4 = '(-inf-129.5)': sick (14.46/1.46)

T3_4 = '(-inf-2.45]' AND
thyroid_surgery = f AND
TT4_measured = t AND
query_hypothyroid = f AND
referral_source = SVHC AND
FTI_3 = '(113.5-inf)': negative (17.21)

T3_4 = '(-inf-2.45]' AND
thyroid_surgery = f AND
TT4_measured = t AND
query_hypothyroid = f AND
T4U_2 = '(0.935-inf)' AND
age_1 = '(32.5-inf)' AND
query_on_thyroxine = f AND
age_2 = '(47.5-inf)' AND
T3_1 = '(-inf-1.47)' AND
age_3 = '(59.5-inf)' AND
on_thyroxine = f AND
referral_source = SVI AND
TT4_4 = '(-inf-131.5]' AND
FTI_3 = '(-inf-113.5]' AND
TSH_4 = '(3.15-inf)': negative (15.53/1.0)

T3_4 = '(-inf-2.45]' AND
thyroid_surgery = f AND
TT4_measured = t AND
query_hypothyroid = f AND
referral_source = other AND
T4U_measured = t AND
T3_2 = '(-inf-1.85]' AND
sex = M AND
age_4 = '(-inf-70.5)': negative (17.61)

thyroid_surgery = f AND
T3_4 = '(-inf-2.45]' AND
TT4_measured = t AND
TT4_4 = '(131.5-inf)' AND
age_1 = '(32.5-inf)' AND
on_thyroxine = f AND
T4U_measured = t AND
T3_3 = '(-inf-2.15]' AND
query_on_thyroxine = f AND
query_hypothyroid = f AND
sex = M AND
age_4 = '(-inf-70.5)': negative (8.12)

thyroid_surgery = f AND
T3_4 = '(-inf-2.45]' AND
TT4_measured = t AND
query_hypothyroid = t AND
sick = f AND
age_2 = '(47.5-inf)' AND
TT4_4 = '(-inf-131.5]' AND
TSH_1 = '(0.295-inf)' AND
T4U_4 = '(-inf-1.115]' AND
referral_source = SVI AND

```

```

FTI_4 = '(-inf-129.5]' AND
FTI_1 = '(89.5-inf)' AND
TSH_4 = '(-inf-3.15]' AND
FTI_3 = '(-inf-113.5)': sick (7.0/1.0)

thyroid_surgery = f AND
T3_4 = '(-inf-2.45]' AND
TT4_measured = t AND
T4U_1 = '(0.855-inf)' AND
query_hyperthyroid = f AND
T3_2 = '(-inf-1.85]' AND
age_1 = '(32.5-inf)' AND
TT4_1 = '(83.5-inf)' AND
psych = f AND
query_on_thyroxine = f AND
referral_source = SVI AND
T4U_measured = t AND
TSH_4 = '(-inf-3.15]' AND
query_hypothyroid = f AND
FTI_1 = '(89.5-inf)' AND
T4U_4 = '(-inf-1.115]' AND
TSH_1 = '(0.295-inf)' AND
T4U_3 = '(-inf-1.015]' AND
sick = f AND
TSH_2 = '(0.995-inf)' AND
TT4_2 = '(97.5-inf)' AND
FTI_3 = '(113.5-inf)': negative (23.0/4.0)

thyroid_surgery = f AND
T3_4 = '(-inf-2.45]' AND
TT4_measured = t AND
age_1 = '(-inf-32.5]' AND
query_hypothyroid = f AND
T4U_2 = '(-inf-0.935]' AND
TT4_3 = '(-inf-111.5)': negative (17.19/1.0)

thyroid_surgery = t: negative (6.14)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
age_1 = '(-inf-32.5]' AND
T4U_1 = '(0.855-inf)' AND
TSH_3 = '(-inf-1.75]' AND
sex = M: sick (4.38/0.55)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
T4U_2 = '(0.935-inf)' AND
age_1 = '(32.5-inf)' AND
sick = f AND
query_on_thyroxine = f AND
age_2 = '(47.5-inf)' AND
FTI_4 = '(129.5-inf)': negative (9.84/1.46)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
T4U_4 = '(1.115-inf)' AND
FTI_1 = '(-inf-89.5)': negative (8.4/0.15)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
referral_source = other AND
age_1 = '(32.5-inf)' AND
T4U_measured = t AND
TT4_1 = '(83.5-inf)' AND
on_thyroxine = f AND
FTI_1 = '(89.5-inf)' AND
sex = F AND
TSH_4 = '(-inf-3.15]' AND
T3_1 = '(-inf-1.47)': negative (12.18/3.0)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
query_on_thyroxine = t AND
TT4_3 = '(111.5-inf)': negative (4.0)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
query_hyperthyroid = t AND
age_2 = '(47.5-inf)' AND
TT4_2 = '(97.5-inf)': sick (3.95/1.0)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND

```

```

T4U_measured = f AND
referral_source = SVI AND
TT4_1 = '(83.5-inf)': sick (6.58/1.68)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
T4U_measured = t AND
TSH_measured = t AND
T4U_1 = '(-inf-0.855]' AND
age_1 = '(32.5-inf)' AND
T3_1 = '(-inf-1.47]' AND
TT4_4 = '(-inf-131.5]' AND
query_hyperthyroid = f AND
FTI_2 = '(101.5-inf)' AND
TSH_4 = '(3.15-inf)' AND
age_4 = '(-inf-70.5]': sick (17.0/1.0)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
T4U_measured = f: negative (10.61/0.98)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
query_hyperthyroid = t AND
age_2 = '(-inf-47.5]': sick (3.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
TT4_1 = '(83.5-inf)' AND
psych = f AND
FTI_1 = '(-inf-89.5]': negative (6.0/1.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
TT4_1 = '(83.5-inf)' AND
psych = f AND
referral_source = other AND
age_1 = '(32.5-inf)' AND
TT4_2 = '(97.5-inf)' AND
on_thyroxine = f AND
sex = M: negative (5.0/1.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
TT4_1 = '(83.5-inf)' AND
psych = f AND
referral_source = other AND
age_1 = '(32.5-inf)' AND
FTI_4 = '(-inf-129.5]': negative (5.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
query_on_thyroxine = f AND
tumor = f AND
TT4_1 = '(-inf-83.5]' AND
sick = f AND
TSH_4 = '(-inf-3.15]' AND
query_hypothyroid = f AND
on_thyroxine = f AND
TSH_2 = '(0.995-inf)' AND
FTI_2 = '(101.5-inf)': sick (8.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
psych = t: negative (7.0/2.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
referral_source = SVHC AND
age_2 = '(47.5-inf)': sick (6.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
TSH_measured = t AND
referral_source = other AND
age_1 = '(32.5-inf)' AND
TSH_2 = '(-inf-0.995]' AND
FTI_2 = '(101.5-inf)' AND
T3_1 = '(-inf-1.47]': negative (3.6/1.0)

```

```

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
TSH_measured = t AND
referral_source = other AND
age_1 = '(32.5-inf)' AND
FTI_2 = '(-inf-101.5]' AND
TSH_1 = '(0.295-inf)' AND
query_hypothyroid = f AND
T4U_1 = '(0.855-inf)': sick (7.0/1.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
TSH_measured = t AND
referral_source = other AND
FTI_2 = '(-inf-101.5)': negative (8.6/1.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
referral_source = other: sick (8.25/1.25)

referral_source = SVI AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
on_thyroxine = t AND
age_4 = '(70.5-inf)': negative (4.0)

referral_source = SVI AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
T3_1 = '(1.47-inf)' AND
TT4_4 = '(-inf-131.5)': negative (4.0)

referral_source = SVI AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
TSH_1 = '(-inf-0.295]' AND
T4U_2 = '(-inf-0.935]' AND
query_on_thyroxine = f AND
FTI_4 = '(-inf-129.5]' AND
age_2 = '(47.5-inf)' AND
FTI_3 = '(113.5-inf)': sick (11.0/1.0)

referral_source = SVI AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
on_thyroxine = f AND
query_on_thyroxine = f AND
T3_3 = '(-inf-2.15]' AND
TT4_1 = '(83.5-inf)' AND
age_2 = '(-inf-47.5]' AND
FTI_2 = '(101.5-inf)' AND
FTI_3 = '(-inf-1.75)': negative (8.0/2.0)

referral_source = SVI AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
on_thyroxine = f AND
query_on_thyroxine = f AND
T3_3 = '(-inf-2.15]' AND
T3_1 = '(-inf-1.47]' AND
TT4_1 = '(83.5-inf)' AND
age_2 = '(47.5-inf)' AND
TSH_3 = '(1.75-inf)' AND
sick = f AND
TSH_4 = '(-inf-3.15]' AND
FTI_2 = '(101.5-inf)' AND
query_hypothyroid = f AND
TT4_3 = '(-inf-111.5]' AND
age_4 = '(-inf-70.5)': negative (8.0/1.0)

referral_source = SVI AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
on_thyroxine = f AND
query_on_thyroxine = f AND

```

```

T3_3 = '(-inf-2.15)' AND
T3_1 = '(1.47-inf)': negative (4.0/1.0)

referral_source = SVHC: negative (3.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_4 = '(-inf-2.45)' AND
query_hyperthyroid = f AND
referral_source = SVI AND
on_thyroxine = f AND
query_on_thyroxine = f AND
T3_1 = '(-inf-1.47)' AND
TT4_4 = '(131.5-inf)' AND
sex = F: negative (3.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_4 = '(-inf-2.45)' AND
query_hyperthyroid = f AND
referral_source = SVI AND
TT4_4 = '(131.5-inf)': sick (6.0/2.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47)' AND
query_hyperthyroid = f AND
referral_source = SVI AND
on_thyroxine = f AND
query_on_thyroxine = f AND
TT4_1 = '(83.5-inf)' AND
age_2 = '(-inf-47.5)': negative (5.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47)' AND
query_hyperthyroid = f AND
referral_source = SVI AND
age_2 = '(-inf-47.5)' AND
T4U_2 = '(-inf-0.935)': sick (11.0/1.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47)' AND
referral_source = SVI AND
query_hyperthyroid = f AND
age_2 = '(47.5-inf)' AND
T4U_1 = '(0.855-inf)' AND
TSH_1 = '(0.295-inf)' AND
TT4_3 = '(111.5-inf)' AND
age_3 = '(59.5-inf)': negative (9.0/1.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47)' AND
query_hyperthyroid = f AND
referral_source = SVI AND
T4U_3 = '(1.015-inf)' AND
age_4 = '(-inf-70.5)' AND
age_3 = '(-inf-59.5)': sick (8.0/2.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47)' AND
query_hyperthyroid = f AND
referral_source = SVI AND
T4U_1 = '(0.855-inf)' AND
age_3 = '(-inf-59.5)': negative (12.0/2.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47)' AND
query_hyperthyroid = f AND
referral_source = SVI AND
TSH_4 = '(3.15-inf)' AND
FTI_3 = '(-inf-113.5)' AND
TT4_2 = '(-inf-97.5)' AND
T4U_2 = '(-inf-0.935)' AND
FTI_1 = '(-inf-89.5)': negative (22.0/6.0)

```

```

T4U_measured = t AND
referral_source = SVI AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
TT4_1 = '(-inf-83.5]' AND
query_hypothyroid = f AND
sex = M AND
TSH_4 = '(-inf-3.15]': sick (20.47/4.47)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
TT4_2 = '(97.5-inf)' AND
age_3 = '(59.5-inf)' AND
query_hypothyroid = f AND
TSH_2 = '(0.995-inf)' AND
T4U_3 = '(-inf-1.015]' AND
sex = F AND
FTI_4 = '(-inf-129.5]': sick (6.56)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
referral_source = SVI AND
query_hyperthyroid = f AND
query_hypothyroid = t AND
TT4_1 = '(-inf-83.5]': sick (4.0)

referral_source = SVI AND
age_1 = '(32.5-inf)' AND
TT4_4 = '(-inf-131.5]' AND
tumor = f AND
query_hyperthyroid = f AND
query_hypothyroid = f AND
FTI_1 = '(-inf-89.5]' AND
age_3 = '(59.5-inf)': negative (7.0/1.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
referral_source = SVI AND
query_hyperthyroid = f AND
query_hypothyroid = f AND
FTI_1 = '(89.5-inf)' AND
FTI_4 = '(129.5-inf)' AND
age_3 = '(59.5-inf)' AND
age_4 = '(70.5-inf)' AND
sick = f AND
TSH_4 = '(-inf-3.15]' AND
TSH_1 = '(0.295-inf)': sick (9.0/3.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
referral_source = SVI AND
query_hyperthyroid = f AND
query_hypothyroid = f AND
FTI_1 = '(89.5-inf)' AND
age_4 = '(70.5-inf)' AND
T4U_3 = '(-inf-1.015]' AND
sick = f AND
sex = M AND
T4U_1 = '(0.855-inf)': negative (10.0/2.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
query_hypothyroid = f AND
FTI_1 = '(89.5-inf)' AND
FTI_3 = '(113.5-inf)' AND
T4U_1 = '(-inf-0.855]' AND
age_4 = '(-inf-70.5]' AND
TSH_1 = '(0.295-inf)': negative (5.98/2.44)

T4U_measured = t AND

```

```

age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
query_hypothyroid = f AND
age_3 = '(59.5-inf)' AND
FTI_3 = '(113.5-inf)' AND
age_4 = '(70.5-inf)' AND
sick = f AND
TSH_4 = '(-inf-3.15]': negative (10.0/3.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
query_hypothyroid = f AND
FTI_4 = '(129.5-inf)' AND
age_4 = '(-inf-70.5]': sick (4.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
query_hypothyroid = f AND
sick = t AND
age_4 = '(70.5-inf)': negative (4.0)

T4U_measured = t AND
TT4_3 = '(-inf-111.5]' AND
query_hyperthyroid = f AND
sex = M AND
query_hypothyroid = f AND
FTI_2 = '(101.5-inf)': sick (9.26/2.26)

T4U_measured = t AND
TT4_3 = '(111.5-inf)': sick (7.0)

query_hyperthyroid = f AND
age_4 = '(-inf-70.5]' AND
TT4_2 = '(-inf-97.5]' AND
sex = F AND
FTI_2 = '(-inf-101.5]': sick (9.81/2.81)

age_4 = '(-inf-70.5]': negative (14.6/1.0)

sex = F AND
FTI_3 = '(-inf-113.5]' AND
TT4_2 = '(-inf-97.5]' AND
TSH_2 = '(-inf-0.995]': sick (4.25/1.01)

sex = F AND
T4U_3 = '(-inf-1.015]' AND
FTI_3 = '(-inf-113.5]' AND
TT4_1 = '(83.5-inf)' AND
TSH_4 = '(-inf-3.15]': negative (4.01)

T4U_1 = '(0.855-inf)': sick (8.02/0.02)

TT4_1 = '(-inf-83.5]': sick (6.78/2.01)

: negative (3.01)

Number of Rules : 75

Time taken to build model: 0.31 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3549          94.088 %
Incorrectly Classified Instances   223           5.912 %
Kappa statistic                   0.4561
Mean absolute error               0.0662
Root mean squared error           0.2146
Relative absolute error            57.453 %
Root relative squared error       89.5079 %
Total Number of Instances         3772

==== Detailed Accuracy By Class ====

```

TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
---------	---------	-----------	--------	-----------	-----	----------	----------	-------

```

0.972    0.541    0.965    0.972    0.969    0.457    0.861    0.975    negative
0.459    0.028    0.520    0.459    0.487    0.457    0.861    0.442    sick
Weighted Avg. 0.941    0.510    0.938    0.941    0.939    0.457    0.861    0.943

==== Confusion Matrix ====
a      b    <-- classified as
3443   98 |     a = negative
125    106 |     b = sick

```

## Cross Validation 10 folds

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B5-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  48
              age_1
              age_2
              age_3
              age_4
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              TSH_2
              TSH_3
              TSH_4
              T3_measured
              T3_1
              T3_2
              T3_3
              T3_4
              TT4_measured
              TT4_1
              TT4_2
              TT4_3
              TT4_4
              T4U_measured
              T4U_1
              T4U_2
              T4U_3
              T4U_4
              FTI_measured
              FTI_1
              FTI_2
              FTI_3
              FTI_4
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   10-fold cross-validation

==== Classifier model (full training set) ===

PART decision list
-----
T3_1 = '(1.47-inf)' AND
FTI_4 = '(-inf-129.5]': negative (2417.11/4.41)

referral_source = other AND
T3_measured = f: negative (265.4/0.2)

T3_1 = '(1.47-inf)' AND
TSH_1 = '(0.295-inf)': negative (230.71/0.17)

referral_source = other AND
sick = f AND

```

```

age_4 = '(-inf-70.5]' AND
T4U_4 = '(-inf-1.115]' AND
TSH_measured = t AND
T4U_measured = t AND
FTI_2 = '(101.5-inf)' AND
TSH_4 = '(-inf-3.15]' AND
on_thyroxine = t: negative (54.0)

referral_source = other AND
T4U_3 = '(1.015-inf)': negative (94.38/1.0)

T3_2 = '(1.85-inf)' AND
TT4_4 = '(131.5-inf)' AND
on_thyroxine = f AND
referral_source = other AND
T3_3 = '(2.15-inf)': negative (43.09/2.0)

T3_measured = t AND
T3_4 = '(2.45-inf)' AND
T4U_1 = '(0.855-inf)': negative (24.96)

T3_measured = t AND
thyroid_surgery = f AND
T3_2 = '(1.85-inf)' AND
TT4_4 = '(-inf-131.5)': negative (25.09)

T3_measured = t AND
T3_4 = '(-inf-2.45)' AND
thyroid_surgery = f AND
referral_source = other AND
T4U_measured = t AND
age_4 = '(-inf-70.5]' AND
TT4_4 = '(-inf-131.5]' AND
on_thyroxine = f AND
T4U_1 = '(0.855-inf)': negative (39.0)

T3_measured = f: negative (17.78/0.22)

T3_4 = '(-inf-2.45)' AND
thyroid_surgery = f AND
T3_1 = '(1.47-inf)' AND
on_thyroxine = f AND
TT4_4 = '(131.5-inf)' AND
age_1 = '(32.5-inf)' AND
sex = F AND
T4U_2 = '(0.935-inf)': negative (12.57/2.73)

T3_4 = '(-inf-2.45)' AND
thyroid_surgery = f AND
T3_1 = '(1.47-inf)' AND
on_thyroxine = t: negative (11.53)

T3_4 = '(-inf-2.45)' AND
thyroid_surgery = f AND
TT4_measured = t AND
sick = t AND
query_hypothyroid = f AND
T4U_3 = '(-inf-1.015]' AND
T4U_1 = '(-inf-0.855]' AND
FTI_4 = '(-inf-129.5)': sick (14.46/1.46)

T3_4 = '(-inf-2.45)' AND
thyroid_surgery = f AND
TT4_measured = t AND
query_hypothyroid = f AND
referral_source = SVHC AND
FTI_3 = '(113.5-inf)': negative (17.21)

T3_4 = '(-inf-2.45)' AND
thyroid_surgery = f AND
TT4_measured = t AND
query_hypothyroid = f AND
T4U_2 = '(0.935-inf)' AND
age_1 = '(32.5-inf)' AND
query_on_thyroxine = f AND
age_2 = '(47.5-inf)' AND
T3_1 = '(-inf-1.47)' AND
age_3 = '(59.5-inf)' AND
on_thyroxine = f AND
referral_source = SVI AND
TT4_4 = '(-inf-131.5)' AND
FTI_3 = '(-inf-113.5)' AND
TSH_4 = '(3.15-inf)': negative (15.53/1.0)

T3_4 = '(-inf-2.45)' AND
thyroid_surgery = f AND
TT4_measured = t AND

```

```

query_hypothyroid = f AND
referral_source = other AND
T4U_measured = t AND
T3_2 = '(-inf-1.85]' AND
sex = M AND
age_4 = '(-inf-70.5]': negative (17.61)

thyroid_surgery = f AND
T3_4 = '(-inf-2.45]' AND
TT4_measured = t AND
TT4_4 = '(131.5-inf)' AND
age_1 = '(32.5-inf)' AND
on_thyroxine = f AND
T4U_measured = t AND
T3_3 = '(-inf-2.15]' AND
query_on_thyroxine = f AND
query_hypothyroid = f AND
sex = M AND
age_4 = '(-inf-70.5]': negative (8.12)

thyroid_surgery = f AND
T3_4 = '(-inf-2.45]' AND
TT4_measured = t AND
query_hypothyroid = t AND
sick = f AND
age_2 = '(47.5-inf)' AND
TT4_4 = '(-inf-131.5]' AND
TSH_1 = '(0.295-inf)' AND
T4U_4 = '(-inf-1.115]' AND
referral_source = SVI AND
FTI_4 = '(-inf-129.5]' AND
FTI_1 = '(89.5-inf)' AND
TSH_4 = '(-inf-3.15]' AND
FTI_3 = '(-inf-113.5]': sick (7.0/1.0)

thyroid_surgery = f AND
T3_4 = '(-inf-2.45]' AND
TT4_measured = t AND
T4U_1 = '(0.855-inf)' AND
query_hyperthyroid = f AND
T3_2 = '(-inf-1.85]' AND
age_1 = '(32.5-inf)' AND
TT4_1 = '(83.5-inf)' AND
psych = f AND
query_on_thyroxine = f AND
referral_source = SVI AND
T4U_measured = t AND
TSH_4 = '(-inf-3.15]' AND
query_hypothyroid = f AND
FTI_1 = '(89.5-inf)' AND
T4U_4 = '(-inf-1.115]' AND
TSH_1 = '(0.295-inf)' AND
T4U_3 = '(-inf-1.015]' AND
sick = f AND
TSH_2 = '(0.995-inf)' AND
TT4_2 = '(97.5-inf)' AND
FTI_3 = '(113.5-inf)': negative (23.0/4.0)

thyroid_surgery = f AND
T3_4 = '(-inf-2.45]' AND
TT4_measured = t AND
age_1 = '(-inf-32.5]' AND
query_hypothyroid = f AND
T4U_2 = '(-inf-0.935]' AND
TT4_3 = '(-inf-111.5]': negative (17.19/1.0)

thyroid_surgery = t: negative (6.14)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
age_1 = '(-inf-32.5]' AND
T4U_1 = '(0.855-inf)' AND
TSH_3 = '(-inf-1.75]' AND
sex = M: sick (4.38/0.55)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
T4U_2 = '(0.935-inf)' AND
age_1 = '(32.5-inf)' AND
sick = f AND
query_on_thyroxine = f AND
age_2 = '(47.5-inf)' AND
FTI_4 = '(129.5-inf)': negative (9.84/1.46)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
T4U_4 = '(1.115-inf)' AND

```

```

FTI_1 = '(-inf-89.5]': negative (8.4/0.15)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
referral_source = other AND
age_1 = '(32.5-inf)' AND
T4U_measured = t AND
TT4_1 = '(83.5-inf)' AND
on_thyroxine = f AND
FTI_1 = '(89.5-inf)' AND
sex = F AND
TSH_4 = '(-inf-3.15]' AND
T3_1 = '(-inf-1.47]': negative (12.18/3.0)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
query_on_thyroxine = t AND
TT4_3 = '(111.5-inf)': negative (4.0)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
query_hyperthyroid = t AND
age_2 = '(47.5-inf)' AND
TT4_2 = '(97.5-inf)': sick (3.95/1.0)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
T4U_measured = f AND
referral_source = SVI AND
TT4_1 = '(83.5-inf)': sick (6.58/1.68)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
T4U_measured = t AND
TSH_measured = t AND
T4U_1 = '(-inf-0.855]' AND
age_1 = '(32.5-inf)' AND
T3_1 = '(-inf-1.47]' AND
TT4_4 = '(-inf-131.5]' AND
query_hyperthyroid = f AND
FTI_2 = '(101.5-inf)' AND
TSH_4 = '(3.15-inf)' AND
age_4 = '(-inf-70.5]': sick (17.0/1.0)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
T4U_measured = f: negative (10.61/0.98)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
query_hyperthyroid = t AND
age_2 = '(-inf-47.5)': sick (3.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
TT4_1 = '(83.5-inf)' AND
psych = f AND
FTI_1 = '(-inf-89.5]': negative (6.0/1.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
TT4_1 = '(83.5-inf)' AND
psych = f AND
referral_source = other AND
age_1 = '(32.5-inf)' AND
TT4_2 = '(97.5-inf)' AND
on_thyroxine = f AND
sex = M: negative (5.0/1.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
TT4_1 = '(83.5-inf)' AND
psych = f AND
referral_source = other AND
age_1 = '(32.5-inf)' AND
FTI_4 = '(-inf-129.5]': negative (5.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND

```

```

query_on_thyroxine = f AND
tumor = f AND
TT4_1 = '(-inf-83.5]' AND
sick = f AND
TSH_4 = '(-inf-3.15]' AND
query_hypothyroid = f AND
on_thyroxine = f AND
TSH_2 = '(0.995-inf)' AND
FTI_2 = '(101.5-inf)': sick (8.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
psych = t: negative (7.0/2.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
referral_source = SVHC AND
age_2 = '(47.5-inf)': sick (6.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
TSH_measured = t AND
referral_source = other AND
age_1 = '(32.5-inf)' AND
TSH_2 = '(-inf-0.995]' AND
FTI_2 = '(101.5-inf)' AND
T3_1 = '(-inf-1.47)': negative (3.6/1.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
TSH_measured = t AND
referral_source = other AND
age_1 = '(32.5-inf)' AND
FTI_2 = '(-inf-101.5]' AND
TSH_1 = '(0.295-inf)' AND
query_hypothyroid = f AND
T4U_1 = '(0.855-inf)': sick (7.0/1.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
TSH_measured = t AND
referral_source = other AND
FTI_2 = '(-inf-101.5)': negative (8.6/1.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
referral_source = other: sick (8.25/1.25)

referral_source = SVI AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
on_thyroxine = t AND
age_4 = '(70.5-inf)': negative (4.0)

referral_source = SVI AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
T3_1 = '(1.47-inf)' AND
T3_4 = '(-inf-131.5)': negative (4.0)

referral_source = SVI AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
TSH_1 = '(-inf-0.295]' AND
T4U_2 = '(-inf-0.935]' AND
query_on_thyroxine = f AND
FTI_4 = '(-inf-129.5]' AND
age_2 = '(47.5-inf)' AND
FTI_3 = '(113.5-inf)': sick (11.0/1.0)

referral_source = SVI AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
on_thyroxine = f AND
query_on_thyroxine = f AND
T3_3 = '(-inf-2.15]' AND
T4_1 = '(83.5-inf)' AND

```

```

age_2 = '(-inf-47.5]' AND
FTI_2 = '(101.5-inf)' AND
TSH_3 = '(-inf-1.75]': negative (8.0/2.0)

referral_source = SVI AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
on_thyroxine = f AND
query_on_thyroxine = f AND
T3_3 = '(-inf-2.15]' AND
T3_1 = '(-inf-1.47]' AND
TT4_1 = '(83.5-inf)' AND
age_2 = '(47.5-inf)' AND
TSH_3 = '(1.75-inf)' AND
sick = f AND
TSH_4 = '(-inf-3.15]' AND
FTI_2 = '(101.5-inf)' AND
query_hypothyroid = f AND
TT4_3 = '(-inf-111.5]' AND
age_4 = '(-inf-70.5]': negative (8.0/1.0)

referral_source = SVI AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
on_thyroxine = f AND
query_on_thyroxine = f AND
T3_3 = '(-inf-2.15]' AND
T3_1 = '(1.47-inf)': negative (4.0/1.0)

referral_source = SVHC: negative (3.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
on_thyroxine = f AND
query_on_thyroxine = f AND
T3_1 = '(-inf-1.47]' AND
TT4_4 = '(131.5-inf)' AND
sex = F: negative (3.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
TT4_4 = '(131.5-inf)': sick (6.0/2.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
on_thyroxine = f AND
query_on_thyroxine = f AND
TT4_1 = '(83.5-inf)' AND
age_2 = '(-inf-47.5]': negative (5.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
age_2 = '(-inf-47.5]' AND
T4U_2 = '(-inf-0.935]': sick (11.0/1.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
referral_source = SVI AND
query_hyperthyroid = f AND
age_2 = '(47.5-inf)' AND
T4U_1 = '(0.855-inf)' AND
TSH_1 = '(0.295-inf)' AND
TT4_3 = '(111.5-inf)' AND
age_3 = '(59.5-inf)': negative (9.0/1.0)

```

```

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
T4U_3 = '(1.015-inf)' AND
age_4 = '(-inf-70.5]' AND
age_3 = '(-inf-59.5]': sick (8.0/2.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
T4U_1 = '(0.855-inf)' AND
age_3 = '(-inf-59.5]': negative (12.0/2.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
TSH_4 = '(3.15-inf)' AND
FTI_3 = '(-inf-113.5]' AND
TT4_2 = '(-inf-97.5]' AND
T4U_2 = '(-inf-0.935]' AND
FTI_1 = '(-inf-89.5]': negative (22.0/6.0)

T4U_measured = t AND
referral_source = SVI AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
TT4_1 = '(-inf-83.5]' AND
query_hypothyroid = f AND
sex = M AND
TSH_4 = '(-inf-3.15]': sick (20.47/4.47)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
TT4_2 = '(97.5-inf)' AND
age_3 = '(59.5-inf)' AND
query_hypothyroid = f AND
TSH_2 = '(0.995-inf)' AND
T4U_3 = '(-inf-1.015]' AND
sex = F AND
FTI_4 = '(-inf-129.5]': sick (6.56)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
referral_source = SVI AND
query_hyperthyroid = f AND
query_hypothyroid = t AND
TT4_1 = '(-inf-83.5]': sick (4.0)

referral_source = SVI AND
age_1 = '(32.5-inf)' AND
TT4_4 = '(-inf-131.5]' AND
tumor = f AND
query_hyperthyroid = f AND
query_hypothyroid = f AND
FTI_1 = '(-inf-89.5]' AND
age_3 = '(59.5-inf)': negative (7.0/1.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
referral_source = SVI AND
query_hyperthyroid = f AND
query_hypothyroid = f AND
FTI_1 = '(89.5-inf)' AND
FTI_4 = '(129.5-inf)' AND
age_3 = '(59.5-inf)' AND
age_4 = '(70.5-inf)' AND
sick = f AND
TSH_4 = '(-inf-3.15]': sick (8.0/2.0)

```

```

TSH_1 = '(0.295-inf)': sick (9.0/3.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
referral_source = SVI AND
query_hyperthyroid = f AND
query_hypothyroid = f AND
FTI_1 = '(89.5-inf)' AND
age_4 = '(70.5-inf)' AND
T4U_3 = '(-inf-1.015]' AND
sick = f AND
sex = M AND
T4U_1 = '(0.855-inf)': negative (10.0/2.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
query_hypothyroid = f AND
FTI_1 = '(89.5-inf)' AND
FTI_3 = '(113.5-inf)' AND
T4U_1 = '(-inf-0.855]' AND
age_4 = '(-inf-70.5]' AND
TSH_1 = '(0.295-inf)': negative (5.98/2.44)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
query_hypothyroid = f AND
age_3 = '(59.5-inf)' AND
FTI_3 = '(113.5-inf)' AND
age_4 = '(70.5-inf)' AND
sick = f AND
TSH_4 = '(-inf-3.15)': negative (10.0/3.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
query_hypothyroid = f AND
FTI_4 = '(129.5-inf)' AND
age_4 = '(-inf-70.5]': sick (4.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
query_hypothyroid = f AND
sick = t AND
age_4 = '(70.5-inf)': negative (4.0)

T4U_measured = t AND
TT4_3 = '(-inf-111.5]' AND
query_hyperthyroid = f AND
sex = M AND
query_hypothyroid = f AND
FTI_2 = '(101.5-inf)': sick (9.26/2.26)

T4U_measured = t AND
TT4_3 = '(111.5-inf)': sick (7.0)

query_hyperthyroid = f AND
age_4 = '(-inf-70.5]' AND
TT4_2 = '(-inf-97.5]' AND
sex = F AND
FTI_2 = '(-inf-101.5]': sick (9.81/2.81)

age_4 = '(-inf-70.5)': negative (14.6/1.0)

sex = F AND
FTI_3 = '(-inf-113.5]' AND
TT4_2 = '(-inf-97.5]' AND
TSH_2 = '(-inf-0.995]': sick (4.25/1.01)

sex = F AND
T4U_3 = '(-inf-1.015]' AND

```

```

FTI_3 = '(-inf-113.5]' AND
TT4_1 = '(83.5-inf)' AND
TSH_4 = '(-inf-3.15]': negative (4.01)

T4U_1 = '(0.855-inf)': sick (8.02/0.02)

TT4_1 = '(-inf-83.5]': sick (6.78/2.01)

: negative (3.01)

Number of Rules : 75

Time taken to build model: 0.3 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3531          93.6108 %
Incorrectly Classified Instances   241           6.3892 %
Kappa statistic                   0.4046
Mean absolute error               0.0681
Root mean squared error          0.2199
Relative absolute error          59.1221 %
Root relative squared error     91.7194 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC    ROC Area  PRC Area  Class
0.971      0.593     0.962     0.971     0.966     0.406    0.894     0.982  negative
0.407      0.029     0.475     0.407     0.438     0.406    0.894     0.436   sick
Weighted Avg.  0.936     0.559     0.932     0.936     0.934     0.406    0.894     0.948

==== Confusion Matrix ====

      a     b  <-- classified as
3437  104 |     a = negative
 137   94 |     b = sick

```

## Cross Validation 15 folds

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B5-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  48
              age_1
              age_2
              age_3
              age_4
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              t131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              TSH_2
              TSH_3
              TSH_4
              T3_measured
              T3_1
              T3_2
              T3_3
              T3_4
              TT4_measured
              TT4_1
              TT4_2
              TT4_3
              TT4_4
              T4U_measured
              T4U_1

```

```

T4U_2
T4U_3
T4U_4
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
TBG_measured
TBG
referral_source
Class
Test mode: 15-fold cross-validation

==== Classifier model (full training set) ====

PART decision list
-----
T3_1 = '(1.47-inf)' AND
FTI_4 = '(-inf-129.5]': negative (2417.11/4.41)

referral_source = other AND
T3_measured = f: negative (265.4/0.2)

T3_1 = '(1.47-inf)' AND
TSH_1 = '(0.295-inf)': negative (230.71/0.17)

referral_source = other AND
sick = f AND
age_4 = '(-inf-70.5]' AND
T4U_4 = '(-inf-1.115]' AND
TSH_measured = t AND
T4U_measured = t AND
FTI_2 = '(101.5-inf)' AND
TSH_4 = '(-inf-3.15]' AND
on_thyroxine = t: negative (54.0)

referral_source = other AND
T4U_3 = '(1.015-inf)': negative (94.38/1.0)

T3_2 = '(1.85-inf)' AND
TT4_4 = '(131.5-inf)' AND
on_thyroxine = f AND
referral_source = other AND
T3_3 = '(2.15-inf)': negative (43.09/2.0)

T3_measured = t AND
T3_4 = '(2.45-inf)' AND
T4U_1 = '(0.855-inf)': negative (24.96)

T3_measured = t AND
thyroid_surgery = f AND
T3_2 = '(1.85-inf)' AND
T4_4 = '(-inf-131.5)': negative (25.09)

T3_measured = t AND
T3_4 = '(-inf-2.45)' AND
thyroid_surgery = f AND
referral_source = other AND
T4U_measured = t AND
age_4 = '(-inf-70.5]' AND
TT4_4 = '(-inf-131.5]' AND
on_thyroxine = f AND
T4U_1 = '(0.855-inf)': negative (39.0)

T3_measured = f: negative (17.78/0.22)

T3_4 = '(-inf-2.45)' AND
thyroid_surgery = f AND
T3_1 = '(1.47-inf)' AND
on_thyroxine = f AND
TT4_4 = '(131.5-inf)' AND
age_1 = '(32.5-inf)' AND
sex = F AND
T4U_2 = '(0.935-inf)': negative (12.57/2.73)

T3_4 = '(-inf-2.45)' AND
thyroid_surgery = f AND
T3_1 = '(1.47-inf)' AND
on_thyroxine = t: negative (11.53)

T3_4 = '(-inf-2.45)' AND
thyroid_surgery = f AND
TT4_measured = t AND
sick = t AND
query_hypothyroid = f AND

```

```

T4U_3 = '(-inf-1.015]' AND
T4U_1 = '(-inf-0.855]' AND
FTI_4 = '(-inf-129.5]': sick (14.46/1.46)

T3_4 = '(-inf-2.45]' AND
thyroid_surgery = f AND
TT4_measured = t AND
query_hypothyroid = f AND
referral_source = SVHC AND
FTI_3 = '(113.5-inf)': negative (17.21)

T3_4 = '(-inf-2.45]' AND
thyroid_surgery = f AND
TT4_measured = t AND
query_hypothyroid = f AND
T4U_2 = '(0.935-inf)' AND
age_1 = '(32.5-inf)' AND
query_on_thyroxine = f AND
age_2 = '(47.5-inf)' AND
T3_1 = '(-inf-1.47)' AND
age_3 = '(59.5-inf)' AND
on_thyroxine = f AND
referral_source = SVI AND
TT4_4 = '(-inf-131.5]' AND
FTI_3 = '(-inf-113.5]' AND
TSH_4 = '(3.15-inf)': negative (15.53/1.0)

T3_4 = '(-inf-2.45]' AND
thyroid_surgery = f AND
TT4_measured = t AND
query_hypothyroid = f AND
referral_source = other AND
T4U_measured = t AND
T3_2 = '(-inf-1.85]' AND
sex = M AND
age_4 = '(-inf-70.5]': negative (17.61)

thyroid_surgery = f AND
T3_4 = '(-inf-2.45]' AND
TT4_measured = t AND
TT4_4 = '(131.5-inf)' AND
age_1 = '(32.5-inf)' AND
on_thyroxine = f AND
T4U_measured = t AND
T3_3 = '(-inf-2.15]' AND
query_on_thyroxine = f AND
query_hypothyroid = f AND
sex = M AND
age_4 = '(-inf-70.5]': negative (8.12)

thyroid_surgery = f AND
T3_4 = '(-inf-2.45]' AND
TT4_measured = t AND
query_hypothyroid = t AND
sick = f AND
age_2 = '(47.5-inf)' AND
TT4_4 = '(-inf-131.5]' AND
TSH_1 = '(0.295-inf)' AND
T4U_4 = '(-inf-1.115]' AND
referral_source = SVI AND
FTI_4 = '(-inf-129.5]' AND
FTI_1 = '(89.5-inf)' AND
TSH_4 = '(-inf-3.15]' AND
FTI_3 = '(-inf-113.5]': sick (7.0/1.0)

thyroid_surgery = f AND
T3_4 = '(-inf-2.45]' AND
TT4_measured = t AND
T4U_1 = '(0.855-inf)' AND
query_hyperthyroid = f AND
T3_2 = '(-inf-1.85]' AND
age_1 = '(32.5-inf)' AND
TT4_1 = '(83.5-inf)' AND
psych = f AND
query_on_thyroxine = f AND
referral_source = SVI AND
T4U_measured = t AND
TSH_4 = '(-inf-3.15]' AND
query_hypothyroid = f AND
FTI_1 = '(89.5-inf)' AND
T4U_4 = '(-inf-1.115]' AND
TSH_1 = '(0.295-inf)' AND
T4U_3 = '(-inf-1.015]' AND
sick = f AND
TSH_2 = '(0.995-inf)' AND
TT4_2 = '(97.5-inf)' AND
FTI_3 = '(113.5-inf)': negative (23.0/4.0)

```

```

thyroid_surgery = f AND
T3_4 = '(-inf-2.45]' AND
TT4_measured = t AND
age_1 = '(-inf-32.5]' AND
query_hypothyroid = f AND
T4U_2 = '(-inf-0.935]' AND
TT4_3 = '(-inf-111.5]': negative (17.19/1.0)

thyroid_surgery = t: negative (6.14)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
age_1 = '(-inf-32.5]' AND
T4U_1 = '(0.855-inf)' AND
TSH_3 = '(-inf-1.75]' AND
sex = M: sick (4.38/0.55)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
T4U_2 = '(0.935-inf)' AND
age_1 = '(32.5-inf)' AND
sick = f AND
query_on_thyroxine = f AND
age_2 = '(47.5-inf)' AND
FTI_4 = '(129.5-inf)': negative (9.84/1.46)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
T4U_4 = '(1.115-inf)' AND
FTI_1 = '(-inf-89.5]': negative (8.4/0.15)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
referral_source = other AND
age_1 = '(32.5-inf)' AND
T4U_measured = t AND
TT4_1 = '(83.5-inf)' AND
on_thyroxine = f AND
FTI_1 = '(89.5-inf)' AND
sex = F AND
TSH_4 = '(-inf-3.15]' AND
T3_1 = '(-inf-1.47]': negative (12.18/3.0)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
query_on_thyroxine = t AND
TT4_3 = '(111.5-inf)': negative (4.0)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
query_hyperthyroid = t AND
age_2 = '(47.5-inf)' AND
T4U_2 = '(97.5-inf)': sick (3.95/1.0)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
T4U_measured = f AND
referral_source = SVI AND
TT4_1 = '(83.5-inf)': sick (6.58/1.68)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
T4U_measured = t AND
TSH_measured = t AND
T4U_1 = '(-inf-0.855]' AND
age_1 = '(32.5-inf)' AND
T3_1 = '(-inf-1.47]': AND
TT4_4 = '(-inf-131.5]' AND
query_hyperthyroid = f AND
FTI_2 = '(101.5-inf)' AND
TSH_4 = '(3.15-inf)' AND
age_4 = '(-inf-70.5]': sick (17.0/1.0)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
T4U_measured = f: negative (10.61/0.98)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
query_hyperthyroid = t AND
age_2 = '(-inf-47.5]': sick (3.0)

```

```

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
TT4_1 = '(83.5-inf)' AND
psych = f AND
FTI_1 = '(-inf-89.5]': negative (6.0/1.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
TT4_1 = '(83.5-inf)' AND
psych = f AND
referral_source = other AND
age_1 = '(32.5-inf)' AND
TT4_2 = '(97.5-inf)' AND
on_thyroxine = f AND
sex = M: negative (5.0/1.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
TT4_1 = '(83.5-inf)' AND
psych = f AND
referral_source = other AND
age_1 = '(32.5-inf)' AND
FTI_4 = '(-inf-129.5]': negative (5.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
query_on_thyroxine = f AND
tumor = f AND
TT4_1 = '(-inf-83.5]' AND
sick = f AND
TSH_4 = '(-inf-3.15]' AND
query_hypothyroid = f AND
on_thyroxine = f AND
TSH_2 = '(0.995-inf)' AND
FTI_2 = '(101.5-inf)': sick (8.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
psych = t: negative (7.0/2.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
referral_source = SVHC AND
age_2 = '(47.5-inf)': sick (6.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
TSH_measured = t AND
referral_source = other AND
age_1 = '(32.5-inf)' AND
TSH_2 = '(-inf-0.995]' AND
FTI_2 = '(101.5-inf)' AND
T3_1 = '(-inf-1.47]': negative (3.6/1.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
TSH_measured = t AND
referral_source = other AND
age_1 = '(32.5-inf)' AND
FTI_2 = '(-inf-101.5]' AND
TSH_1 = '(0.295-inf)' AND
query_hypothyroid = f AND
T4U_1 = '(0.855-inf)': sick (7.0/1.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
TSH_measured = t AND
referral_source = other AND
FTI_2 = '(-inf-101.5]': negative (8.6/1.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
referral_source = other: sick (8.25/1.25)

referral_source = SVI AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND

```

```

query_hyperthyroid = f AND
on_thyroxine = t AND
age_4 = '(70.5-inf)': negative (4.0)

referral_source = SVI AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
T3_1 = '(1.47-inf)' AND
TT4_4 = '(-inf-131.5]': negative (4.0)

referral_source = SVI AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
TSH_1 = '(-inf-0.295]' AND
T4U_2 = '(-inf-0.935]' AND
query_on_thyroxine = f AND
FTI_4 = '(-inf-129.5]' AND
age_2 = '(47.5-inf)' AND
FTI_3 = '(113.5-inf)': sick (11.0/1.0)

referral_source = SVI AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
on_thyroxine = f AND
query_on_thyroxine = f AND
T3_3 = '(-inf-2.15]' AND
TT4_1 = '(83.5-inf)' AND
age_2 = '(-inf-47.5]' AND
FTI_2 = '(101.5-inf)' AND
TSH_3 = '(-inf-1.75]': negative (8.0/2.0)

referral_source = SVI AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
on_thyroxine = f AND
query_on_thyroxine = f AND
T3_3 = '(-inf-2.15]' AND
T3_1 = '(-inf-1.47]' AND
TT4_1 = '(83.5-inf)' AND
age_2 = '(47.5-inf)' AND
TSH_3 = '(1.75-inf)' AND
sick = f AND
TSH_4 = '(-inf-3.15]' AND
FTI_2 = '(101.5-inf)' AND
query_hypothyroid = f AND
TT4_3 = '(-inf-111.5]' AND
age_4 = '(-inf-70.5]': negative (8.0/1.0)

referral_source = SVI AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
on_thyroxine = f AND
query_on_thyroxine = f AND
T3_3 = '(-inf-2.15]' AND
T3_1 = '(1.47-inf)': negative (4.0/1.0)

referral_source = SVHC: negative (3.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
on_thyroxine = f AND
query_on_thyroxine = f AND
T3_1 = '(-inf-1.47]' AND
TT4_4 = '(131.5-inf)' AND
sex = F: negative (3.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
TT4_4 = '(131.5-inf)': sick (6.0/2.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND

```

```

tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
on_thyroxine = f AND
query_on_thyroxine = f AND
TT4_1 = '(83.5-inf)' AND
age_2 = '(-inf-47.5]': negative (5.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
age_2 = '(-inf-47.5]' AND
T4U_2 = '(-inf-0.935]': sick (11.0/1.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
referral_source = SVI AND
query_hyperthyroid = f AND
age_2 = '(47.5-inf)' AND
T4U_1 = '(0.855-inf)' AND
TSH_1 = '(0.295-inf)' AND
TT4_3 = '(111.5-inf)' AND
age_3 = '(59.5-inf)': negative (9.0/1.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
T4U_3 = '(1.015-inf)' AND
age_4 = '(-inf-70.5]' AND
age_3 = '(-inf-59.5]': sick (8.0/2.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
T4U_1 = '(0.855-inf)' AND
age_3 = '(-inf-59.5]': negative (12.0/2.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
TSH_4 = '(3.15-inf)' AND
FTI_3 = '(-inf-113.5]' AND
TT4_2 = '(-inf-97.5]' AND
T4U_2 = '(-inf-0.935]' AND
T4U_1 = '(-inf-89.5]': negative (22.0/6.0)

T4U_measured = t AND
referral_source = SVI AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
TT4_1 = '(-inf-83.5]' AND
query_hypothyroid = f AND
sex = M AND
TSH_4 = '(-inf-3.15]': sick (20.47/4.47)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
TT4_2 = '(97.5-inf)' AND
age_3 = '(59.5-inf)' AND
query_hypothyroid = f AND
TSH_2 = '(0.995-inf)' AND
T4U_3 = '(-inf-1.015]' AND
sex = F AND
FTI_4 = '(-inf-129.5]': sick (6.56)

T4U_measured = t AND

```

```

age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
referral_source = SVI AND
query_hyperthyroid = f AND
query_hypothyroid = t AND
TT4_1 = '(-inf-83.5]': sick (4.0)

referral_source = SVI AND
age_1 = '(32.5-inf)' AND
TT4_4 = '(-inf-131.5]' AND
tumor = f AND
query_hyperthyroid = f AND
query_hypothyroid = f AND
FTI_1 = '(-inf-89.5]' AND
age_3 = '(59.5-inf)': negative (7.0/1.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
referral_source = SVI AND
query_hyperthyroid = f AND
query_hypothyroid = f AND
FTI_1 = '(89.5-inf)' AND
FTI_4 = '(129.5-inf)' AND
age_3 = '(59.5-inf)' AND
age_4 = '(70.5-inf)' AND
sick = f AND
TSH_4 = '(-inf-3.15]' AND
TSH_1 = '(0.295-inf)': sick (9.0/3.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
referral_source = SVI AND
query_hyperthyroid = f AND
query_hypothyroid = f AND
FTI_1 = '(89.5-inf)' AND
age_4 = '(70.5-inf)' AND
T4U_3 = '(-inf-1.015]' AND
sick = f AND
sex = M AND
T4U_1 = '(0.855-inf)': negative (10.0/2.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
query_hypothyroid = f AND
FTI_1 = '(89.5-inf)' AND
FTI_3 = '(113.5-inf)' AND
T4U_1 = '(-inf-0.855]' AND
age_4 = '(-inf-70.5]' AND
TSH_1 = '(0.295-inf)': negative (5.98/2.44)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
query_hypothyroid = f AND
age_3 = '(59.5-inf)' AND
FTI_3 = '(113.5-inf)' AND
age_4 = '(70.5-inf)' AND
sick = f AND
TSH_4 = '(-inf-3.15]': negative (10.0/3.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
query_hypothyroid = f AND
FTI_4 = '(129.5-inf)' AND
age_4 = '(-inf-70.5]': sick (4.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND

```

```

referral_source = SVI AND
query_hypothyroid = f AND
sick = t AND
age_4 = '(70.5-inf)': negative (4.0)

T4U_measured = t AND
TT4_3 = '(-inf-111.5]' AND
query_hyperthyroid = f AND
sex = M AND
query_hypothyroid = f AND
FTI_2 = '(101.5-inf)': sick (9.26/2.26)

T4U_measured = t AND
TT4_3 = '(111.5-inf)': sick (7.0)

query_hyperthyroid = f AND
age_4 = '(-inf-70.5]' AND
TT4_2 = '(-inf-97.5]' AND
sex = F AND
FTI_2 = '(-inf-101.5)': sick (9.81/2.81)

age_4 = '(-inf-70.5)': negative (14.6/1.0)

sex = F AND
FTI_3 = '(-inf-113.5]' AND
TT4_2 = '(-inf-97.5]' AND
TSH_2 = '(-inf-0.995)': sick (4.25/1.01)

sex = F AND
T4U_3 = '(-inf-1.015]' AND
FTI_3 = '(-inf-113.5]' AND
TT4_1 = '(83.5-inf)' AND
TSH_4 = '(-inf-3.15)': negative (4.01)

T4U_1 = '(0.855-inf)': sick (8.02/0.02)

TT4_1 = '(-inf-83.5)': sick (6.78/2.01)

: negative (3.01)

Number of Rules : 75

Time taken to build model: 0.29 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3557           94.3001 %
Incorrectly Classified Instances   215            5.6999 %
Kappa statistic                   0.4618
Mean absolute error               0.0633
Root mean squared error          0.2087
Relative absolute error           54.9423 %
Root relative squared error      87.0326 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC     ROC Area  PRC Area  Class
      0.975    0.550    0.965    0.975    0.970    0.464    0.887    0.980    negative
      0.450    0.025    0.542    0.450    0.492    0.464    0.887    0.503    sick
Weighted Avg.      0.943    0.518    0.939    0.943    0.941    0.464    0.887    0.951

==== Confusion Matrix ====

      a      b  <-- classified as
3453   88 |  a = negative
  127  104 |  b = sick

```

## Percentage Split 50%

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B5-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  48
              age_1
              age_2
              age_3
              age_4
              sex
              on_thyroxine

```

```

query_on_thyroxine
on_antithyroid_medication
sick
pregnant
thyroid_surgery
I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH_1
TSH_2
TSH_3
TSH_4
T3_measured
T3_1
T3_2
T3_3
T3_4
TT4_measured
TT4_1
TT4_2
TT4_3
TT4_4
T4U_measured
T4U_1
T4U_2
T4U_3
T4U_4
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
TBG_measured
TBG
referral_source
Class
Test mode: split 50.0% train, remainder test

==== Classifier model (full training set) ===

PART decision list
-----
T3_1 = '(1.47-inf)' AND
FTI_4 = '(-inf-129.5)': negative (2417.11/4.41)

referral_source = other AND
T3_measured = f: negative (265.4/0.2)

T3_1 = '(1.47-inf)' AND
TSH_1 = '(0.295-inf)': negative (230.71/0.17)

referral_source = other AND
sick = f AND
age_4 = '(-inf-70.5]' AND
T4U_4 = '(-inf-1.115]' AND
TSH_measured = t AND
T4U_measured = t AND
FTI_2 = '(101.5-inf)' AND
TSH_4 = '(-inf-3.15]' AND
on_thyroxine = t: negative (54.0)

referral_source = other AND
T4U_3 = '(1.015-inf)': negative (94.38/1.0)

T3_2 = '(1.85-inf)' AND
TT4_4 = '(131.5-inf)' AND
on_thyroxine = f AND
referral_source = other AND
T3_3 = '(2.15-inf)': negative (43.09/2.0)

T3_measured = t AND
T3_4 = '(2.45-inf)' AND
T4U_1 = '(0.855-inf)': negative (24.96)

T3_measured = t AND
thyroid_surgery = f AND
T3_2 = '(1.85-inf)' AND
TT4_4 = '(-inf-131.5)': negative (25.09)

T3_measured = t AND

```

```

T3_4 = '(-inf-2.45]' AND
thyroid_surgery = f AND
referral_source = other AND
T4U_measured = t AND
age_4 = '(-inf-70.5]' AND
TT4_4 = '(-inf-131.5]' AND
on_thyroxine = f AND
T4U_1 = '(0.855-inf)': negative (39.0)

T3_measured = f: negative (17.78/0.22)

T3_4 = '(-inf-2.45]' AND
thyroid_surgery = f AND
T3_1 = '(1.47-inf)' AND
on_thyroxine = f AND
TT4_4 = '(131.5-inf)' AND
age_1 = '(32.5-inf)' AND
sex = F AND
T4U_2 = '(0.935-inf)': negative (12.57/2.73)

T3_4 = '(-inf-2.45]' AND
thyroid_surgery = f AND
T3_1 = '(1.47-inf)' AND
on_thyroxine = t: negative (11.53)

T3_4 = '(-inf-2.45]' AND
thyroid_surgery = f AND
TT4_measured = t AND
sick = t AND
query_hypothyroid = f AND
T4U_3 = '(-inf-1.015]' AND
T4U_1 = '(-inf-0.855]' AND
FTI_4 = '(-inf-129.5)': sick (14.46/1.46)

T3_4 = '(-inf-2.45]' AND
thyroid_surgery = f AND
TT4_measured = t AND
query_hypothyroid = f AND
referral_source = SVHC AND
FTI_3 = '(113.5-inf)': negative (17.21)

T3_4 = '(-inf-2.45]' AND
thyroid_surgery = f AND
TT4_measured = t AND
query_hypothyroid = f AND
T4U_2 = '(0.935-inf)' AND
age_1 = '(32.5-inf)' AND
query_on_thyroxine = f AND
age_2 = '(47.5-inf)' AND
T3_1 = '(-inf-1.47]' AND
age_3 = '(59.5-inf)' AND
on_thyroxine = f AND
referral_source = SVI AND
TT4_4 = '(-inf-131.5]' AND
FTI_3 = '(-inf-113.5]' AND
TSH_4 = '(3.15-inf)': negative (15.53/1.0)

T3_4 = '(-inf-2.45]' AND
thyroid_surgery = f AND
TT4_measured = t AND
query_hypothyroid = f AND
referral_source = other AND
T4U_measured = t AND
T3_2 = '(-inf-1.85]' AND
sex = M AND
age_4 = '(-inf-70.5)': negative (17.61)

thyroid_surgery = f AND
T3_4 = '(-inf-2.45]' AND
TT4_measured = t AND
TT4_4 = '(131.5-inf)' AND
age_1 = '(32.5-inf)' AND
on_thyroxine = f AND
T4U_measured = t AND
T3_3 = '(-inf-2.15]' AND
query_on_thyroxine = f AND
query_hypothyroid = f AND
sex = M AND
age_4 = '(-inf-70.5)': negative (8.12)

thyroid_surgery = f AND
T3_4 = '(-inf-2.45]' AND
TT4_measured = t AND
query_hypothyroid = t AND
sick = f AND
age_2 = '(47.5-inf)' AND
TT4_4 = '(-inf-131.5]' AND

```

```

TSH_1 = '(0.295-inf)' AND
T4U_4 = '(-inf-1.115]' AND
referral_source = SVI AND
FTI_4 = '(-inf-129.5]' AND
FTI_1 = '(89.5-inf)' AND
TSH_4 = '(-inf-3.15]' AND
FTI_3 = '(-inf-113.5]': sick (7.0/1.0)

thyroid_surgery = f AND
T3_4 = '(-inf-2.45]' AND
TT4_measured = t AND
T4U_1 = '(0.855-inf)' AND
query_hyperthyroid = f AND
T3_2 = '(-inf-1.85]' AND
age_1 = '(32.5-inf)' AND
TT4_1 = '(83.5-inf)' AND
psych = f AND
query_on_thyroxine = f AND
referral_source = SVI AND
T4U_measured = t AND
TSH_4 = '(-inf-3.15]' AND
query_hypothyroid = f AND
FTI_1 = '(89.5-inf)' AND
T4U_4 = '(-inf-1.115]' AND
TSH_1 = '(0.295-inf)' AND
T4U_3 = '(-inf-1.015]' AND
sick = f AND
TSH_2 = '(0.995-inf)' AND
TT4_2 = '(97.5-inf)' AND
FTI_3 = '(113.5-inf)': negative (23.0/4.0)

thyroid_surgery = f AND
T3_4 = '(-inf-2.45]' AND
TT4_measured = t AND
age_1 = '(-inf-32.5]' AND
query_hypothyroid = f AND
T4U_2 = '(-inf-0.935]' AND
TT4_3 = '(-inf-111.5]': negative (17.19/1.0)

thyroid_surgery = t: negative (6.14)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
age_1 = '(-inf-32.5]' AND
T4U_1 = '(0.855-inf)' AND
TSH_3 = '(-inf-1.75]' AND
sex = M: sick (4.38/0.55)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
T4U_2 = '(0.935-inf)' AND
age_1 = '(32.5-inf)' AND
sick = f AND
query_on_thyroxine = f AND
age_2 = '(47.5-inf)' AND
FTI_4 = '(129.5-inf)': negative (9.84/1.46)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
T4U_4 = '(1.115-inf)' AND
FTI_1 = '(-inf-89.5]': negative (8.4/0.15)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
referral_source = other AND
age_1 = '(32.5-inf)' AND
T4U_measured = t AND
TT4_1 = '(83.5-inf)' AND
on_thyroxine = f AND
FTI_1 = '(89.5-inf)' AND
sex = F AND
TSH_4 = '(-inf-3.15]' AND
T3_1 = '(-inf-1.47]': negative (12.18/3.0)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
query_on_thyroxine = t AND
TT4_3 = '(111.5-inf)': negative (4.0)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
query_hyperthyroid = t AND
age_2 = '(47.5-inf)' AND
TT4_2 = '(97.5-inf)': sick (3.95/1.0)

```

```

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
T4U_measured = f AND
referral_source = SVI AND
TT4_1 = '(83.5-inf)': sick (6.58/1.68)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
T4U_measured = t AND
TSH_measured = t AND
T4U_1 = '(-inf-0.855]' AND
age_1 = '(32.5-inf)' AND
T3_1 = '(-inf-1.47]' AND
TT4_4 = '(-inf-131.5]' AND
query_hyperthyroid = f AND
FTI_2 = '(101.5-inf)' AND
TSH_4 = '(3.15-inf)' AND
age_4 = '(-inf-70.5)': sick (17.0/1.0)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
T4U_measured = f: negative (10.61/0.98)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
query_hyperthyroid = t AND
age_2 = '(-inf-47.5)': sick (3.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
TT4_1 = '(83.5-inf)' AND
psych = f AND
FTI_1 = '(-inf-89.5)': negative (6.0/1.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
TT4_1 = '(83.5-inf)' AND
psych = f AND
referral_source = other AND
age_1 = '(32.5-inf)' AND
TT4_2 = '(97.5-inf)' AND
on_thyroxine = f AND
sex = M: negative (5.0/1.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
TT4_1 = '(83.5-inf)' AND
psych = f AND
referral_source = other AND
age_1 = '(32.5-inf)' AND
FTI_4 = '(-inf-129.5)': negative (5.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
query_on_thyroxine = f AND
tumor = f AND
TT4_1 = '(-inf-83.5]' AND
sick = f AND
TSH_4 = '(-inf-3.15]' AND
query_hypothyroid = f AND
on_thyroxine = f AND
TSH_2 = '(0.995-inf)' AND
FTI_2 = '(101.5-inf)': sick (8.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
psych = t: negative (7.0/2.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
referral_source = SVHC AND
age_2 = '(47.5-inf)': sick (6.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
TSH_measured = t AND
referral_source = other AND
age_1 = '(32.5-inf)' AND

```

```

TSH_2 = '(-inf-0.995]' AND
FTI_2 = '(101.5-inf)' AND
T3_1 = '(-inf-1.47]': negative (3.6/1.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
TSH_measured = t AND
referral_source = other AND
age_1 = '(32.5-inf)' AND
FTI_2 = '(-inf-101.5]' AND
TSH_1 = '(0.295-inf)' AND
query_hypothyroid = f AND
T4U_1 = '(0.855-inf)': sick (7.0/1.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
TSH_measured = t AND
referral_source = other AND
FTI_2 = '(-inf-101.5]': negative (8.6/1.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
referral_source = other: sick (8.25/1.25)

referral_source = SVI AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
on_thyroxine = t AND
age_4 = '(70.5-inf)': negative (4.0)

referral_source = SVI AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
T3_1 = '(1.47-inf)' AND
TT4_4 = '(-inf-131.5]': negative (4.0)

referral_source = SVI AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
TSH_1 = '(-inf-0.295]' AND
T4U_2 = '(-inf-0.935]' AND
query_on_thyroxine = f AND
FTI_4 = '(-inf-129.5]' AND
age_2 = '(47.5-inf)' AND
FTI_3 = '(113.5-inf)': sick (11.0/1.0)

referral_source = SVI AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
on_thyroxine = f AND
query_on_thyroxine = f AND
T3_3 = '(-inf-2.15]' AND
TT4_1 = '(83.5-inf)' AND
age_2 = '(-inf-47.5]' AND
FTI_2 = '(101.5-inf)' AND
TSH_3 = '(-inf-1.75]': negative (8.0/2.0)

referral_source = SVI AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
on_thyroxine = f AND
query_on_thyroxine = f AND
T3_3 = '(-inf-2.15]' AND
T3_1 = '(-inf-1.47]' AND
TT4_1 = '(83.5-inf)' AND
age_2 = '(47.5-inf)' AND
TSH_3 = '(1.75-inf)' AND
sick = f AND
TSH_4 = '(-inf-3.15]' AND
FTI_2 = '(101.5-inf)' AND
query_hypothyroid = f AND
TT4_3 = '(-inf-111.5]' AND
age_4 = '(-inf-70.5)': negative (8.0/1.0)

referral_source = SVI AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND

```

```

query_hyperthyroid = f AND
on_thyroxine = f AND
query_on_thyroxine = f AND
T3_3 = '(-inf-2.15]' AND
T3_1 = '(1.47-inf)': negative (4.0/1.0)

referral_source = SVHC: negative (3.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
on_thyroxine = f AND
query_on_thyroxine = f AND
T3_1 = '(-inf-1.47]' AND
TT4_4 = '(131.5-inf)' AND
sex = F: negative (3.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
TT4_4 = '(131.5-inf)': sick (6.0/2.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
on_thyroxine = f AND
query_on_thyroxine = f AND
TT4_1 = '(83.5-inf)' AND
age_2 = '(-inf-47.5]': negative (5.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
age_2 = '(-inf-47.5]' AND
T4U_2 = '(-inf-0.935)': sick (11.0/1.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
referral_source = SVI AND
query_hyperthyroid = f AND
age_2 = '(47.5-inf)' AND
T4U_1 = '(0.855-inf)' AND
TSH_1 = '(0.295-inf)' AND
TT4_3 = '(111.5-inf)' AND
age_3 = '(59.5-inf)': negative (9.0/1.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
T4U_3 = '(1.015-inf)' AND
age_4 = '(-inf-70.5]' AND
age_3 = '(-inf-59.5)': sick (8.0/2.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
T4U_1 = '(0.855-inf)' AND
age_3 = '(-inf-59.5)': negative (12.0/2.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
TSH_4 = '(3.15-inf)' AND
FTI_3 = '(-inf-113.5]' AND

```

```

TT4_2 = '(-inf-97.5]' AND
T4U_2 = '(-inf-0.935]' AND
FTI_1 = '(-inf-89.5]': negative (22.0/6.0)

T4U_measured = t AND
referral_source = SVI AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
TT4_1 = '(-inf-83.5]' AND
query_hypothyroid = f AND
sex = M AND
TSH_4 = '(-inf-3.15]': sick (20.47/4.47)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
TT4_2 = '(97.5-inf)' AND
age_3 = '(59.5-inf)' AND
query_hypothyroid = f AND
TSH_2 = '(0.995-inf)' AND
T4U_3 = '(-inf-1.015]' AND
sex = F AND
TSH_4 = '(-inf-129.5]': sick (6.56)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
referral_source = SVI AND
query_hyperthyroid = f AND
query_hypothyroid = t AND
TT4_1 = '(-inf-83.5]': sick (4.0)

referral_source = SVI AND
age_1 = '(32.5-inf)' AND
TT4_4 = '(-inf-131.5]' AND
tumor = f AND
query_hyperthyroid = f AND
query_hypothyroid = f AND
FTI_1 = '(-inf-89.5]' AND
age_3 = '(59.5-inf)': negative (7.0/1.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
referral_source = SVI AND
query_hyperthyroid = f AND
query_hypothyroid = f AND
FTI_1 = '(89.5-inf)' AND
FTI_4 = '(129.5-inf)' AND
age_3 = '(59.5-inf)' AND
age_4 = '(70.5-inf)' AND
sick = f AND
TSH_4 = '(-inf-3.15]' AND
TSH_1 = '(0.295-inf)': sick (9.0/3.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
referral_source = SVI AND
query_hyperthyroid = f AND
query_hypothyroid = f AND
FTI_1 = '(89.5-inf)' AND
age_4 = '(70.5-inf)' AND
T4U_3 = '(-inf-1.015]' AND
sick = f AND
sex = M AND
T4U_1 = '(0.855-inf)': negative (10.0/2.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
query_hypothyroid = f AND
FTI_1 = '(89.5-inf)' AND
FTI_3 = '(113.5-inf)' AND
T4U_1 = '(-inf-0.855]' AND
age_4 = '(-inf-70.5]' AND

```

```

TSH_1 = '(0.295-inf)': negative (5.98/2.44)
T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
query_hypothyroid = f AND
age_3 = '(59.5-inf)' AND
FTI_3 = '(113.5-inf)' AND
age_4 = '(70.5-inf)' AND
sick = f AND
TSH_4 = '(-inf-3.15]': negative (10.0/3.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
query_hypothyroid = f AND
FTI_4 = '(129.5-inf)' AND
age_4 = '(-inf-70.5)': sick (4.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
query_hypothyroid = f AND
sick = t AND
age_4 = '(70.5-inf)': negative (4.0)

T4U_measured = t AND
TT4_3 = '(-inf-111.5]' AND
query_hyperthyroid = f AND
sex = M AND
query_hypothyroid = f AND
FTI_2 = '(101.5-inf)': sick (9.26/2.26)

T4U_measured = t AND
TT4_3 = '(111.5-inf)': sick (7.0)

query_hyperthyroid = f AND
age_4 = '(-inf-70.5]' AND
TT4_2 = '(-inf-97.5]' AND
sex = F AND
FTI_2 = '(-inf-101.5)': sick (9.81/2.81)

age_4 = '(-inf-70.5)': negative (14.6/1.0)

sex = F AND
FTI_3 = '(-inf-113.5]' AND
TT4_2 = '(-inf-97.5]' AND
TSH_2 = '(-inf-0.995)': sick (4.25/1.01)

sex = F AND
T4U_3 = '(-inf-1.015]' AND
FTI_3 = '(-inf-113.5]' AND
TT4_1 = '(83.5-inf)' AND
TSH_4 = '(-inf-3.15]': negative (4.01)

T4U_1 = '(0.855-inf)': sick (8.02/0.02)

TT4_1 = '(-inf-83.5)': sick (6.78/2.01)

: negative (3.01)

Number of Rules : 75

Time taken to build model: 0.29 seconds
==== Evaluation on test split ====
Time taken to test model on test split: 0 seconds
==== Summary ====
Correctly Classified Instances      1772          93.9555 %
Incorrectly Classified Instances   114           6.0445 %
Kappa statistic                   0.4306
Mean absolute error               0.0662
Root mean squared error          0.2155
Relative absolute error           57.4028 %

```

```

Root relative squared error          89.3344 %
Total Number of Instances          1886

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
0.974     0.581     0.962     0.974     0.968     0.433     0.894     0.985   negative
0.419     0.026     0.516     0.419     0.462     0.433     0.894     0.441   sick
Weighted Avg.    0.940     0.547     0.934     0.940     0.937     0.433     0.894     0.951

==== Confusion Matrix ====

     a     b  <-- classified as
1723  46 |     a = negative
     68  49 |     b = sick

```

## Percentage Split 60%

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B5-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  48
              age_1
              age_2
              age_3
              age_4
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              TSH_2
              TSH_3
              TSH_4
              T3_measured
              T3_1
              T3_2
              T3_3
              T3_4
              TT4_measured
              TT4_1
              TT4_2
              TT4_3
              TT4_4
              T4U_measured
              T4U_1
              T4U_2
              T4U_3
              T4U_4
              FTI_measured
              FTI_1
              FTI_2
              FTI_3
              FTI_4
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 60.0% train, remainder test

==== Classifier model (full training set) ===

PART decision list
-----
T3_1 = '(1.47-inf)' AND
FTI_4 = '(-inf-129.5]': negative (2417.11/4.41)

referral_source = other AND
T3_measured = f: negative (265.4/0.2)

```

```

T3_1 = '(1.47-inf)' AND
TSH_1 = '(0.295-inf)': negative (230.71/0.17)

referral_source = other AND
sick = f AND
age_4 = '(-inf-70.5]' AND
T4U_4 = '(-inf-1.115]' AND
TSH_measured = t AND
T4U_measured = t AND
FTI_2 = '(101.5-inf)' AND
TSH_4 = '(-inf-3.15]' AND
on_thyroxine = t: negative (54.0)

referral_source = other AND
T4U_3 = '(1.015-inf)': negative (94.38/1.0)

T3_2 = '(1.85-inf)' AND
TT4_4 = '(131.5-inf)' AND
on_thyroxine = f AND
referral_source = other AND
T3_3 = '(2.15-inf)': negative (43.09/2.0)

T3_measured = t AND
T3_4 = '(2.45-inf)' AND
T4U_1 = '(0.855-inf)': negative (24.96)

T3_measured = t AND
thyroid_surgery = f AND
T3_2 = '(1.85-inf)' AND
TT4_4 = '(-inf-131.5)': negative (25.09)

T3_measured = t AND
T3_4 = '(-inf-2.45]' AND
thyroid_surgery = f AND
referral_source = other AND
T4U_measured = t AND
age_4 = '(-inf-70.5]' AND
TT4_4 = '(-inf-131.5]' AND
on_thyroxine = f AND
T4U_1 = '(0.855-inf)': negative (39.0)

T3_measured = f: negative (17.78/0.22)

T3_4 = '(-inf-2.45)' AND
thyroid_surgery = f AND
T3_1 = '(1.47-inf)' AND
on_thyroxine = f AND
TT4_4 = '(131.5-inf)' AND
age_1 = '(32.5-inf)' AND
sex = F AND
T4U_2 = '(0.935-inf)': negative (12.57/2.73)

T3_4 = '(-inf-2.45]' AND
thyroid_surgery = f AND
T3_1 = '(1.47-inf)' AND
on_thyroxine = t: negative (11.53)

T3_4 = '(-inf-2.45)' AND
thyroid_surgery = f AND
TT4_measured = t AND
sick = t AND
query_hypothyroid = f AND
T4U_3 = '(-inf-1.015]' AND
T4U_1 = '(-inf-0.855]' AND
FTI_4 = '(-inf-129.5)': sick (14.46/1.46)

T3_4 = '(-inf-2.45]' AND
thyroid_surgery = f AND
TT4_measured = t AND
query_hypothyroid = f AND
referral_source = SVHC AND
FTI_3 = '(113.5-inf)': negative (17.21)

T3_4 = '(-inf-2.45)' AND
thyroid_surgery = f AND
TT4_measured = t AND
query_hypothyroid = f AND
T4U_2 = '(0.935-inf)' AND
age_1 = '(32.5-inf)' AND
query_on_thyroxine = f AND
age_2 = '(47.5-inf)' AND
T3_1 = '(-inf-1.47]' AND
age_3 = '(59.5-inf)' AND
on_thyroxine = f AND
referral_source = SVI AND
TT4_4 = '(-inf-131.5]' AND

```

```

FTI_3 = '(-inf-113.5]' AND
TSH_4 = '(3.15-inf)': negative (15.53/1.0)

T3_4 = '(-inf-2.45]' AND
thyroid_surgery = f AND
TT4_measured = t AND
query_hypothyroid = f AND
referral_source = other AND
T4U_measured = t AND
T3_2 = '(-inf-1.85]' AND
sex = M AND
age_4 = '(-inf-70.5]': negative (17.61)

thyroid_surgery = f AND
T3_4 = '(-inf-2.45]' AND
TT4_measured = t AND
TT4_4 = '(131.5-inf)' AND
age_1 = '(32.5-inf)' AND
on_thyroxine = f AND
T4U_measured = t AND
T3_3 = '(-inf-2.15]' AND
query_on_thyroxine = f AND
query_hypothyroid = f AND
sex = M AND
age_4 = '(-inf-70.5]': negative (8.12)

thyroid_surgery = f AND
T3_4 = '(-inf-2.45]' AND
TT4_measured = t AND
query_hypothyroid = t AND
sick = f AND
age_2 = '(47.5-inf)' AND
TT4_4 = '(-inf-131.5]' AND
TSH_1 = '(0.295-inf)' AND
T4U_4 = '(-inf-1.115]' AND
referral_source = SVI AND
FTI_4 = '(-inf-129.5]' AND
FTI_1 = '(89.5-inf)' AND
TSH_4 = '(-inf-3.15]' AND
FTI_3 = '(-inf-113.5]': sick (7.0/1.0)

thyroid_surgery = f AND
T3_4 = '(-inf-2.45]' AND
TT4_measured = t AND
T4U_1 = '(0.855-inf)' AND
query_hyperthyroid = f AND
T3_2 = '(-inf-1.85]' AND
age_1 = '(32.5-inf)' AND
TT4_1 = '(83.5-inf)' AND
psych = f AND
query_on_thyroxine = f AND
referral_source = SVI AND
T4U_measured = t AND
TSH_4 = '(-inf-3.15]' AND
query_hypothyroid = f AND
FTI_1 = '(89.5-inf)' AND
T4U_4 = '(-inf-1.115]' AND
TSH_1 = '(0.295-inf)' AND
T4U_3 = '(-inf-1.015]' AND
sick = f AND
TSH_2 = '(0.995-inf)' AND
TT4_2 = '(97.5-inf)' AND
FTI_3 = '(113.5-inf)': negative (23.0/4.0)

thyroid_surgery = f AND
T3_4 = '(-inf-2.45]' AND
TT4_measured = t AND
age_1 = '(-inf-32.5]' AND
query_hypothyroid = f AND
T4U_2 = '(-inf-0.935]' AND
TT4_3 = '(-inf-111.5]': negative (17.19/1.0)

thyroid_surgery = t: negative (6.14)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
age_1 = '(-inf-32.5]' AND
T4U_1 = '(0.855-inf)' AND
TSH_3 = '(-inf-1.75]' AND
sex = M: sick (4.38/0.55)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
T4U_2 = '(0.935-inf)' AND
age_1 = '(32.5-inf)' AND
sick = f AND
query_on_thyroxine = f AND

```

```

age_2 = '(47.5-inf)' AND
FTI_4 = '(129.5-inf)': negative (9.84/1.46)

TT4_measured = t AND
T3_4 = '(-inf-2.45)' AND
T4U_4 = '(1.115-inf)' AND
FTI_1 = '(-inf-89.5)': negative (8.4/0.15)

TT4_measured = t AND
T3_4 = '(-inf-2.45)' AND
referral_source = other AND
age_1 = '(32.5-inf)' AND
T4U_measured = t AND
TT4_1 = '(83.5-inf)' AND
on_thyroxine = f AND
FTI_1 = '(89.5-inf)' AND
sex = F AND
TSH_4 = '(-inf-3.15]' AND
T3_1 = '(-inf-1.47)': negative (12.18/3.0)

TT4_measured = t AND
T3_4 = '(-inf-2.45)' AND
TSH_measured = t AND
query_on_thyroxine = t AND
TT4_3 = '(111.5-inf)': negative (4.0)

TT4_measured = t AND
T3_4 = '(-inf-2.45)' AND
TSH_measured = t AND
query_hyperthyroid = t AND
age_2 = '(47.5-inf)' AND
TT4_2 = '(97.5-inf)': sick (3.95/1.0)

TT4_measured = t AND
T3_4 = '(-inf-2.45)' AND
TSH_measured = t AND
T4U_measured = f AND
referral_source = SVI AND
TT4_1 = '(83.5-inf)': sick (6.58/1.68)

TT4_measured = t AND
T3_4 = '(-inf-2.45)' AND
T4U_measured = t AND
TSH_measured = t AND
T4U_1 = '(-inf-0.855]' AND
age_1 = '(32.5-inf)' AND
T3_1 = '(-inf-1.47)' AND
TT4_4 = '(-inf-131.5)' AND
query_hyperthyroid = f AND
FTI_2 = '(101.5-inf)' AND
TSH_4 = '(3.15-inf)' AND
age_4 = '(-inf-70.5)': sick (17.0/1.0)

TT4_measured = t AND
T3_4 = '(-inf-2.45)' AND
T4U_measured = f: negative (10.61/0.98)

T4U_measured = t AND
T3_4 = '(-inf-2.45)' AND
TSH_measured = t AND
query_hyperthyroid = t AND
age_2 = '(-inf-47.5)': sick (3.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45)' AND
TSH_measured = t AND
TT4_1 = '(83.5-inf)' AND
psych = f AND
FTI_1 = '(-inf-89.5)': negative (6.0/1.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45)' AND
TSH_measured = t AND
TT4_1 = '(83.5-inf)' AND
psych = f AND
referral_source = other AND
age_1 = '(32.5-inf)' AND
TT4_2 = '(97.5-inf)' AND
on_thyroxine = f AND
sex = M: negative (5.0/1.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45)' AND
TSH_measured = t AND
TT4_1 = '(83.5-inf)' AND
psych = f AND
referral_source = other AND

```

```

age_1 = '(32.5-inf)' AND
FTI_4 = '(-inf-129.5)': negative (5.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45)' AND
TSH_measured = t AND
query_on_thyroxine = f AND
tumor = f AND
TT4_1 = '(-inf-83.5)' AND
sick = f AND
TSH_4 = '(-inf-3.15)' AND
query_hypothyroid = f AND
on_thyroxine = f AND
TSH_2 = '(0.995-inf)' AND
FTI_2 = '(101.5-inf)': sick (8.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45)' AND
TSH_measured = t AND
psych = t: negative (7.0/2.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45)' AND
TSH_measured = t AND
referral_source = SVHC AND
age_2 = '(47.5-inf)': sick (6.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45)' AND
query_hyperthyroid = f AND
TSH_measured = t AND
referral_source = other AND
age_1 = '(32.5-inf)' AND
TSH_2 = '(-inf-0.995)' AND
FTI_2 = '(101.5-inf)' AND
T3_1 = '(-inf-1.47)': negative (3.6/1.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45)' AND
query_hyperthyroid = f AND
TSH_measured = t AND
referral_source = other AND
age_1 = '(32.5-inf)' AND
FTI_2 = '(-inf-101.5)' AND
TSH_1 = '(0.295-inf)' AND
query_hypothyroid = f AND
T4U_1 = '(0.855-inf)': sick (7.0/1.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45)' AND
query_hyperthyroid = f AND
TSH_measured = t AND
referral_source = other AND
FTI_2 = '(-inf-101.5)': negative (8.6/1.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45)' AND
query_hyperthyroid = f AND
referral_source = other: sick (8.25/1.25)

referral_source = SVI AND
tumor = f AND
T3_4 = '(-inf-2.45)' AND
query_hyperthyroid = f AND
on_thyroxine = t AND
age_4 = '(70.5-inf)': negative (4.0)

referral_source = SVI AND
tumor = f AND
T3_4 = '(-inf-2.45)' AND
query_hyperthyroid = f AND
T3_1 = '(1.47-inf)' AND
TT4_4 = '(-inf-131.5)': negative (4.0)

referral_source = SVI AND
tumor = f AND
T3_4 = '(-inf-2.45)' AND
query_hyperthyroid = f AND
TSH_1 = '(-inf-0.295)' AND
T4U_2 = '(-inf-0.935)' AND
query_on_thyroxine = f AND
FTI_4 = '(-inf-129.5)' AND
age_2 = '(47.5-inf)' AND
FTI_3 = '(113.5-inf)': sick (11.0/1.0)

referral_source = SVI AND
tumor = f AND

```

```

T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
on_thyroxine = f AND
query_on_thyroxine = f AND
T3_3 = '(-inf-2.15]' AND
TT4_1 = '(83.5-inf)' AND
age_2 = '(-inf-47.5]' AND
FTI_2 = '(101.5-inf)' AND
TSH_3 = '(-inf-1.75]': negative (8.0/2.0)

referral_source = SVI AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
on_thyroxine = f AND
query_on_thyroxine = f AND
T3_3 = '(-inf-2.15]' AND
T3_1 = '(-inf-1.47]' AND
TT4_1 = '(83.5-inf)' AND
age_2 = '(47.5-inf)' AND
TSH_3 = '(1.75-inf)' AND
sick = f AND
TSH_4 = '(-inf-3.15]' AND
FTI_2 = '(101.5-inf)' AND
query_hypothyroid = f AND
TT4_3 = '(-inf-111.5]' AND
age_4 = '(-inf-70.5]': negative (8.0/1.0)

referral_source = SVI AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
on_thyroxine = f AND
query_on_thyroxine = f AND
T3_3 = '(-inf-2.15]' AND
T3_1 = '(1.47-inf)': negative (4.0/1.0)

referral_source = SVHC: negative (3.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
on_thyroxine = f AND
query_on_thyroxine = f AND
T3_1 = '(-inf-1.47]' AND
TT4_4 = '(131.5-inf)' AND
sex = F: negative (3.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
TT4_4 = '(131.5-inf)': sick (6.0/2.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
on_thyroxine = f AND
query_on_thyroxine = f AND
TT4_1 = '(83.5-inf)' AND
age_2 = '(-inf-47.5]': negative (5.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
age_2 = '(-inf-47.5]' AND
T4U_2 = '(-inf-0.935]': sick (11.0/1.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
referral_source = SVI AND
query_hyperthyroid = f AND

```

```

age_2 = '(47.5-inf)' AND
T4U_1 = '(0.855-inf)' AND
TSH_1 = '(0.295-inf)' AND
TT4_3 = '(111.5-inf)' AND
age_3 = '(59.5-inf)': negative (9.0/1.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
T4U_3 = '(1.015-inf)' AND
age_4 = '(-inf-70.5]' AND
age_3 = '(-inf-59.5]': sick (8.0/2.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
T4U_1 = '(0.855-inf)' AND
age_3 = '(-inf-59.5)': negative (12.0/2.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
TSH_4 = '(3.15-inf)' AND
FTI_3 = '(-inf-113.5]' AND
TT4_2 = '(-inf-97.5]' AND
T4U_2 = '(-inf-0.935]' AND
FTI_1 = '(-inf-89.5)': negative (22.0/6.0)

T4U_measured = t AND
referral_source = SVI AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
TT4_1 = '(-inf-83.5]' AND
query_hypothyroid = f AND
sex = M AND
TSH_4 = '(-inf-3.15]': sick (20.47/4.47)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
TT4_2 = '(97.5-inf)' AND
age_3 = '(59.5-inf)' AND
query_hypothyroid = f AND
TSH_2 = '(0.995-inf)' AND
T4U_3 = '(-inf-1.015]' AND
sex = F AND
FTI_4 = '(-inf-129.5)': sick (6.56)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
referral_source = SVI AND
query_hyperthyroid = f AND
query_hypothyroid = t AND
TT4_1 = '(-inf-83.5)': sick (4.0)

referral_source = SVI AND
age_1 = '(32.5-inf)' AND
TT4_4 = '(-inf-131.5]' AND
tumor = f AND
query_hyperthyroid = f AND
query_hypothyroid = f AND
FTI_1 = '(-inf-89.5)' AND
age_3 = '(59.5-inf)': negative (7.0/1.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
referral_source = SVI AND
query_hyperthyroid = f AND
query_hypothyroid = f AND

```

```

FTI_1 = '(89.5-inf)' AND
FTI_4 = '(129.5-inf)' AND
age_3 = '(59.5-inf)' AND
age_4 = '(70.5-inf)' AND
sick = f AND
TSH_4 = '(-inf-3.15]' AND
TSH_1 = '(0.295-inf)': sick (9.0/3.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
referral_source = SVI AND
query_hyperthyroid = f AND
query_hypothyroid = f AND
FTI_1 = '(89.5-inf)' AND
age_4 = '(70.5-inf)' AND
T4U_3 = '(-inf-1.015]' AND
sick = f AND
sex = M AND
T4U_1 = '(0.855-inf)': negative (10.0/2.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
query_hypothyroid = f AND
FTI_1 = '(89.5-inf)' AND
FTI_3 = '(113.5-inf)' AND
T4U_1 = '(-inf-0.855)' AND
age_4 = '(-inf-70.5]' AND
TSH_1 = '(0.295-inf)': negative (5.98/2.44)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
query_hypothyroid = f AND
age_3 = '(59.5-inf)' AND
FTI_3 = '(113.5-inf)' AND
age_4 = '(70.5-inf)' AND
sick = f AND
TSH_4 = '(-inf-3.15)': negative (10.0/3.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
query_hypothyroid = f AND
FTI_4 = '(129.5-inf)' AND
age_4 = '(-inf-70.5)': sick (4.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
query_hypothyroid = f AND
sick = t AND
age_4 = '(70.5-inf)': negative (4.0)

T4U_measured = t AND
TT4_3 = '(-inf-111.5)' AND
query_hyperthyroid = f AND
sex = M AND
query_hypothyroid = f AND
FTI_2 = '(101.5-inf)': sick (9.26/2.26)

T4U_measured = t AND
TT4_3 = '(111.5-inf)': sick (7.0)

query_hyperthyroid = f AND
age_4 = '(-inf-70.5]' AND
TT4_2 = '(-inf-97.5]' AND
sex = F AND
FTI_2 = '(-inf-101.5)': sick (9.81/2.81)

age_4 = '(-inf-70.5)': negative (14.6/1.0)

sex = F AND

```

```

FTI_3 = '(-inf-113.5]' AND
TT4_2 = '(-inf-97.5]' AND
TSH_2 = '(-inf-0.995]': sick (4.25/1.01)

sex = F AND
T4U_3 = '(-inf-1.015]' AND
FTI_3 = '(-inf-113.5]' AND
TT4_1 = '(83.5-inf)' AND
TSH_4 = '(-inf-3.15]': negative (4.01)

T4U_1 = '(0.855-inf)': sick (8.02/0.02)

TT4_1 = '(-inf-83.5]': sick (6.78/2.01)

: negative (3.01)

Number of Rules : 75

Time taken to build model: 0.34 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

==== Summary ===

Correctly Classified Instances      1419          94.0358 %
Incorrectly Classified Instances   90           5.9642 %
Kappa statistic                   0.4264
Mean absolute error               0.0664
Root mean squared error          0.2202
Relative absolute error          57.8641 %
Root relative squared error     93.4613 %
Total Number of Instances        1509

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC    ROC Area   PRC Area   Class
      0.973     0.573     0.964     0.973     0.968     0.428     0.807     0.969     negative
      0.427     0.027     0.494     0.427     0.458     0.428     0.807     0.370     sick
Weighted Avg.       0.940     0.541     0.937     0.940     0.938     0.428     0.807     0.934

==== Confusion Matrix ===

      a     b  <-- classified as
1381  39 |     a = negative
      51  38 |     b = sick

```

## Percentage Split 66%

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B5-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  48
             age_1
             age_2
             age_3
             age_4
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             I131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary
             psych
             TSH_measured
             TSH_1
             TSH_2
             TSH_3
             TSH_4
             T3_measured
             T3_1
             T3_2

```

```

T3_3
T3_4
TT4_measured
TT4_1
TT4_2
TT4_3
TT4_4
T4U_measured
T4U_1
T4U_2
T4U_3
T4U_4
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
TBG_measured
TBG
referral_source
Class
Test mode: split 66.0% train, remainder test

==== Classifier model (full training set) ===

PART decision list
-----
T3_1 = '(1.47-inf)' AND
FTI_4 = '(-inf-129.5)': negative (2417.11/4.41)

referral_source = other AND
T3_measured = f: negative (265.4/0.2)

T3_1 = '(1.47-inf)' AND
TSH_1 = '(0.295-inf)': negative (230.71/0.17)

referral_source = other AND
sick = f AND
age_4 = '(-inf-70.5]' AND
T4U_4 = '(-inf-1.115]' AND
TSH_measured = t AND
T4U_measured = t AND
FTI_2 = '(101.5-inf)' AND
TSH_4 = '(-inf-3.15]' AND
on_thyroxine = t: negative (54.0)

referral_source = other AND
T4U_3 = '(1.015-inf)': negative (94.38/1.0)

T3_2 = '(1.85-inf)' AND
TT4_4 = '(131.5-inf)' AND
on_thyroxine = f AND
referral_source = other AND
T3_3 = '(2.15-inf)': negative (43.09/2.0)

T3_measured = t AND
T3_4 = '(2.45-inf)' AND
T4U_1 = '(0.855-inf)': negative (24.96)

T3_measured = t AND
thyroid_surgery = f AND
T3_2 = '(1.85-inf)' AND
TT4_4 = '(-inf-131.5)': negative (25.09)

T3_measured = t AND
T3_4 = '(-inf-2.45)' AND
thyroid_surgery = f AND
referral_source = other AND
T4U_measured = t AND
age_4 = '(-inf-70.5]' AND
TT4_4 = '(-inf-131.5)' AND
on_thyroxine = f AND
T4U_1 = '(0.855-inf)': negative (39.0)

T3_measured = f: negative (17.78/0.22)

T3_4 = '(-inf-2.45)' AND
thyroid_surgery = f AND
T3_1 = '(1.47-inf)' AND
on_thyroxine = f AND
TT4_4 = '(131.5-inf)' AND
age_1 = '(32.5-inf)' AND
sex = F AND
T4U_2 = '(0.935-inf)': negative (12.57/2.73)

T3_4 = '(-inf-2.45)' AND

```

```

thyroid_surgery = f AND
T3_1 = '(1.47-inf)' AND
on_thyroxine = t: negative (11.53)

T3_4 = '(-inf-2.45)' AND
thyroid_surgery = f AND
TT4_measured = t AND
sick = t AND
query_hypothyroid = f AND
T4U_3 = '(-inf-1.015)' AND
T4U_1 = '(-inf-0.855)' AND
FTI_4 = '(-inf-129.5)': sick (14.46/1.46)

T3_4 = '(-inf-2.45)' AND
thyroid_surgery = f AND
TT4_measured = t AND
query_hypothyroid = f AND
referral_source = SVHC AND
FTI_3 = '(113.5-inf)': negative (17.21)

T3_4 = '(-inf-2.45)' AND
thyroid_surgery = f AND
TT4_measured = t AND
query_hypothyroid = f AND
T4U_2 = '(0.935-inf)' AND
age_1 = '(32.5-inf)' AND
query_on_thyroxine = f AND
age_2 = '(47.5-inf)' AND
T3_1 = '(-inf-1.47)' AND
age_3 = '(59.5-inf)' AND
on_thyroxine = f AND
referral_source = SVI AND
TT4_4 = '(-inf-131.5)' AND
FTI_3 = '(-inf-113.5)' AND
TSH_4 = '(3.15-inf)': negative (15.53/1.0)

T3_4 = '(-inf-2.45)' AND
thyroid_surgery = f AND
TT4_measured = t AND
query_hypothyroid = f AND
referral_source = other AND
T4U_measured = t AND
T3_2 = '(-inf-1.85)' AND
sex = M AND
age_4 = '(-inf-70.5)': negative (17.61)

thyroid_surgery = f AND
T3_4 = '(-inf-2.45)' AND
TT4_measured = t AND
query_hypothyroid = f AND
age_1 = '(32.5-inf)' AND
on_thyroxine = f AND
T4U_measured = t AND
T3_3 = '(-inf-2.15)' AND
query_on_thyroxine = f AND
query_hypothyroid = f AND
sex = M AND
age_4 = '(-inf-70.5)': negative (8.12)

thyroid_surgery = f AND
T3_4 = '(-inf-2.45)' AND
TT4_measured = t AND
query_hypothyroid = t AND
sick = f AND
age_2 = '(47.5-inf)' AND
TT4_4 = '(-inf-131.5)' AND
TSH_1 = '(0.295-inf)' AND
T4U_4 = '(-inf-1.115)' AND
referral_source = SVI AND
FTI_4 = '(-inf-129.5)' AND
FTI_1 = '(89.5-inf)' AND
TSH_4 = '(-inf-3.15)' AND
FTI_3 = '(-inf-113.5)': sick (7.0/1.0)

thyroid_surgery = f AND
T3_4 = '(-inf-2.45)' AND
TT4_measured = t AND
T4U_1 = '(0.855-inf)' AND
query_hyperthyroid = f AND
T3_2 = '(-inf-1.85)' AND
age_1 = '(32.5-inf)' AND
TT4_1 = '(83.5-inf)' AND
psych = f AND
query_on_thyroxine = f AND
referral_source = SVI AND
T4U_measured = t AND
TSH_4 = '(-inf-3.15)' AND

```

```

query_hypothyroid = f AND
FTI_1 = '(89.5-inf)' AND
T4U_4 = '(-inf-1.115]' AND
TSH_1 = '(0.295-inf)' AND
T4U_3 = '(-inf-1.015]' AND
sick = f AND
TSH_2 = '(0.995-inf)' AND
TT4_2 = '(97.5-inf)' AND
FTI_3 = '(113.5-inf)': negative (23.0/4.0)

thyroid_surgery = f AND
T3_4 = '(-inf-2.45]' AND
TT4_measured = t AND
age_1 = '(-inf-32.5]' AND
query_hypothyroid = f AND
T4U_2 = '(-inf-0.935]' AND
TT4_3 = '(-inf-111.5)': negative (17.19/1.0)

thyroid_surgery = t: negative (6.14)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
age_1 = '(-inf-32.5]' AND
T4U_1 = '(0.855-inf)' AND
TSH_3 = '(-inf-1.75]' AND
sex = M: sick (4.38/0.55)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
T4U_2 = '(0.935-inf)' AND
age_1 = '(32.5-inf)' AND
sick = f AND
query_on_thyroxine = f AND
age_2 = '(47.5-inf)' AND
FTI_4 = '(129.5-inf)': negative (9.84/1.46)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
T4U_4 = '(1.115-inf)' AND
FTI_1 = '(-inf-89.5)': negative (8.4/0.15)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
referral_source = other AND
age_1 = '(32.5-inf)' AND
T4U_measured = t AND
TT4_1 = '(83.5-inf)' AND
on_thyroxine = f AND
FTI_1 = '(89.5-inf)' AND
sex = F AND
TSH_4 = '(-inf-3.15]' AND
T3_1 = '(-inf-1.47)': negative (12.18/3.0)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
query_on_thyroxine = t AND
TT4_3 = '(111.5-inf)': negative (4.0)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
query_hyperthyroid = t AND
age_2 = '(47.5-inf)' AND
TT4_2 = '(97.5-inf)': sick (3.95/1.0)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
T4U_measured = f AND
referral_source = SVI AND
TT4_1 = '(83.5-inf)': sick (6.58/1.68)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
T4U_measured = t AND
TSH_measured = t AND
T4U_1 = '(-inf-0.855]' AND
age_1 = '(32.5-inf)' AND
T3_1 = '(-inf-1.47]' AND
TT4_4 = '(-inf-131.5]' AND
query_hyperthyroid = f AND
FTI_2 = '(101.5-inf)' AND
TSH_4 = '(3.15-inf)' AND
age_4 = '(-inf-70.5)': sick (17.0/1.0)

TT4_measured = t AND

```

```

T3_4 = '(-inf-2.45]' AND
T4U_measured = f: negative (10.61/0.98)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
query_hyperthyroid = t AND
age_2 = '(-inf-47.5]': sick (3.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
TT4_1 = '(83.5-inf)' AND
psych = f AND
FTI_1 = '(-inf-89.5]': negative (6.0/1.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
TT4_1 = '(83.5-inf)' AND
psych = f AND
referral_source = other AND
age_1 = '(32.5-inf)' AND
TT4_2 = '(97.5-inf)' AND
on_thyroxine = f AND
sex = M: negative (5.0/1.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
TT4_1 = '(83.5-inf)' AND
psych = f AND
referral_source = other AND
age_1 = '(32.5-inf)' AND
FTI_4 = '(-inf-129.5]': negative (5.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
query_on_thyroxine = f AND
tumor = f AND
TT4_1 = '(-inf-83.5]' AND
sick = f AND
TSH_4 = '(-inf-3.15]' AND
query_hypothyroid = f AND
on_thyroxine = f AND
TSH_2 = '(0.995-inf)' AND
TTI_2 = '(101.5-inf)': sick (8.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
psych = t: negative (7.0/2.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
referral_source = SVHC AND
age_2 = '(47.5-inf)': sick (6.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
TSH_measured = t AND
referral_source = other AND
age_1 = '(32.5-inf)' AND
TSH_2 = '(-inf-0.995]' AND
FTI_2 = '(101.5-inf)' AND
T3_1 = '(-inf-1.47]': negative (3.6/1.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
TSH_measured = t AND
referral_source = other AND
age_1 = '(32.5-inf)' AND
FTI_2 = '(-inf-101.5]' AND
TSH_1 = '(0.295-inf)' AND
query_hypothyroid = f AND
T4U_1 = '(0.855-inf)': sick (7.0/1.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
TSH_measured = t AND
referral_source = other AND
FTI_2 = '(-inf-101.5]': negative (8.6/1.0)

```

```

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
referral_source = other: sick (8.25/1.25)

referral_source = SVI AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
on_thyroxine = t AND
on_thyroxine = '(-inf-70.5]': negative (4.0)

referral_source = SVI AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
T3_1 = '(1.47-inf)' AND
TT4_4 = '(-inf-131.5]': negative (4.0)

referral_source = SVI AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
TSH_1 = '(-inf-0.295]' AND
T4U_2 = '(-inf-0.935]' AND
query_on_thyroxine = f AND
FTI_4 = '(-inf-129.5]' AND
age_2 = '(47.5-inf)' AND
FTI_3 = '(113.5-inf)': sick (11.0/1.0)

referral_source = SVI AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
on_thyroxine = f AND
query_on_thyroxine = f AND
T3_3 = '(-inf-2.15]' AND
TT4_1 = '(83.5-inf)' AND
age_2 = '(-inf-47.5]' AND
FTI_2 = '(101.5-inf)' AND
TSH_3 = '(-inf-1.75]': negative (8.0/2.0)

referral_source = SVI AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
on_thyroxine = f AND
query_on_thyroxine = f AND
T3_3 = '(-inf-2.15]' AND
T3_1 = '(-inf-1.47]' AND
TT4_1 = '(83.5-inf)' AND
age_2 = '(47.5-inf)' AND
TSH_3 = '(1.75-inf)' AND
sick = f AND
TSH_4 = '(-inf-3.15]' AND
FTI_2 = '(101.5-inf)' AND
query_hypothyroid = f AND
TT4_3 = '(-inf-111.5]' AND
age_4 = '(-inf-70.5]': negative (8.0/1.0)

referral_source = SVI AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
on_thyroxine = f AND
query_on_thyroxine = f AND
T3_3 = '(-inf-2.15]' AND
T3_1 = '(1.47-inf)': negative (4.0/1.0)

referral_source = SVHC: negative (3.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
on_thyroxine = f AND
query_on_thyroxine = f AND
T3_1 = '(-inf-1.47]' AND
TT4_4 = '(131.5-inf)' AND
sex = F: negative (3.0)

T4U_measured = t AND

```

```

age_1 = '(32.5-inf)' AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
TT4_4 = '(131.5-inf)': sick (6.0/2.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
on_thyroxine = f AND
query_on_thyroxine = f AND
TT4_1 = '(83.5-inf)' AND
age_2 = '(-inf-47.5]': negative (5.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
age_2 = '(-inf-47.5]': negative (5.0)
T4U_2 = '(-inf-0.935)': sick (11.0/1.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
referral_source = SVI AND
query_hyperthyroid = f AND
age_2 = '(47.5-inf)' AND
T4U_1 = '(0.855-inf)' AND
TSH_1 = '(0.295-inf)' AND
TT4_3 = '(111.5-inf)' AND
age_3 = '(59.5-inf)': negative (9.0/1.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
T4U_3 = '(1.015-inf)' AND
age_4 = '(-inf-70.5]' AND
age_3 = '(-inf-59.5]': sick (8.0/2.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
T4U_1 = '(0.855-inf)' AND
age_3 = '(-inf-59.5]': negative (12.0/2.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
TSH_4 = '(3.15-inf)' AND
FTI_3 = '(-inf-113.5]' AND
TT4_2 = '(-inf-97.5]' AND
T4U_2 = '(-inf-0.935)' AND
FTI_1 = '(-inf-89.5]': negative (22.0/6.0)

T4U_measured = t AND
referral_source = SVI AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
TT4_1 = '(-inf-83.5]' AND
query_hypothyroid = f AND
sex = M AND
TSH_4 = '(-inf-3.15]': sick (20.47/4.47)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND

```

```

TT4_2 = '(97.5-inf)' AND
age_3 = '(59.5-inf)' AND
query_hypothyroid = f AND
TSH_2 = '(0.995-inf)' AND
T4U_3 = '(-inf-1.015]' AND
sex = F AND
FTI_4 = '(-inf-129.5]': sick (6.56)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
referral_source = SVI AND
query_hyperthyroid = f AND
query_hypothyroid = t AND
TT4_1 = '(-inf-83.5]': sick (4.0)

referral_source = SVI AND
age_1 = '(32.5-inf)' AND
TT4_4 = '(-inf-131.5]' AND
tumor = f AND
query_hyperthyroid = f AND
query_hypothyroid = f AND
FTI_1 = '(-inf-89.5]' AND
age_3 = '(59.5-inf)': negative (7.0/1.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
referral_source = SVI AND
query_hyperthyroid = f AND
query_hypothyroid = f AND
FTI_1 = '(89.5-inf)' AND
FTI_4 = '(129.5-inf)' AND
age_3 = '(59.5-inf)' AND
age_4 = '(70.5-inf)' AND
sick = f AND
TSH_4 = '(-inf-3.15]' AND
TSH_1 = '(0.295-inf)': sick (9.0/3.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
referral_source = SVI AND
query_hyperthyroid = f AND
query_hypothyroid = f AND
FTI_1 = '(89.5-inf)' AND
age_4 = '(70.5-inf)' AND
T4U_3 = '(-inf-1.015]' AND
sick = f AND
sex = M AND
T4U_1 = '(0.855-inf)': negative (10.0/2.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
query_hypothyroid = f AND
FTI_1 = '(89.5-inf)' AND
FTI_3 = '(113.5-inf)' AND
T4U_1 = '(-inf-0.855]' AND
age_4 = '(-inf-70.5]' AND
TSH_1 = '(0.295-inf)': negative (5.98/2.44)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
query_hypothyroid = f AND
age_3 = '(59.5-inf)' AND
FTI_3 = '(113.5-inf)' AND
age_4 = '(70.5-inf)' AND
sick = f AND
TSH_4 = '(-inf-3.15]': negative (10.0/3.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND

```

```

query_hypothyroid = f AND
FTI_4 = '(129.5-inf)' AND
age_4 = '(-inf-70.5]': sick (4.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
query_hypothyroid = f AND
sick = t AND
age_4 = '(70.5-inf)': negative (4.0)

T4U_measured = t AND
TT4_3 = '(-inf-111.5]' AND
query_hyperthyroid = f AND
sex = M AND
query_hypothyroid = f AND
FTI_2 = '(101.5-inf)': sick (9.26/2.26)

T4U_measured = t AND
TT4_3 = '(111.5-inf)': sick (7.0)

query_hyperthyroid = f AND
age_4 = '(-inf-70.5]' AND
TT4_2 = '(-inf-97.5]' AND
sex = F AND
FTI_2 = '(-inf-101.5)': sick (9.81/2.81)

age_4 = '(-inf-70.5)': negative (14.6/1.0)

sex = F AND
FTI_3 = '(-inf-113.5]' AND
TT4_2 = '(-inf-97.5]' AND
TSH_2 = '(-inf-0.995]': sick (4.25/1.01)

sex = F AND
T4U_3 = '(-inf-1.015]' AND
FTI_3 = '(-inf-113.5]' AND
TT4_1 = '(83.5-inf)' AND
TSH_4 = '(-inf-3.15)': negative (4.01)

T4U_1 = '(0.855-inf)': sick (8.02/0.02)

TT4_1 = '(-inf-83.5)': sick (6.78/2.01)

: negative (3.01)

Number of Rules : 75

Time taken to build model: 0.32 seconds

==== Evaluation on test split ====

Time taken to test model on test split: 0 seconds

==== Summary ====

    Correctly Classified Instances      1196           93.2917 %
    Incorrectly Classified Instances     86            6.7083 %
    Kappa statistic                      0.3754
    Mean absolute error                  0.0723
    Root mean squared error              0.2253
    Relative absolute error              63.3786 %
    Root relative squared error          95.982  %
    Total Number of Instances            1282

==== Detailed Accuracy By Class ====

    TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
        0.966    0.600    0.963     0.966    0.964     0.376    0.882     0.984    negative
        0.400    0.034    0.423     0.400    0.411     0.376    0.882     0.402    sick
Weighted Avg.    0.933    0.567    0.931     0.933    0.932     0.376    0.882     0.950

==== Confusion Matrix ====

    a      b  <-- classified as
1166   41 |   a = negative
    45   30 |   b = sick

```

## Leave one out fold

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B5-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  48
              age_1
              age_2
              age_3
              age_4
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              t131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              TSH_2
              TSH_3
              TSH_4
              T3_measured
              T3_1
              T3_2
              T3_3
              T3_4
              TT4_measured
              TT4_1
              TT4_2
              TT4_3
              TT4_4
              T4U_measured
              T4U_1
              T4U_2
              T4U_3
              T4U_4
              FTI_measured
              FTI_1
              FTI_2
              FTI_3
              FTI_4
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   3772-fold cross-validation

==== Classifier model (full training set) ====

PART decision list
-----
T3_1 = '(1.47-inf)' AND
FTI_4 = '(-inf-129.5]': negative (2417.11/4.41)

referral_source = other AND
T3_measured = f: negative (265.4/0.2)

T3_1 = '(1.47-inf)' AND
TSH_1 = '(0.295-inf)': negative (230.71/0.17)

referral_source = other AND
sick = f AND
age_4 = '(-inf-70.5]' AND
T4U_4 = '(-inf-1.115]' AND
TSH_measured = t AND
T4U_measured = t AND
FTI_2 = '(101.5-inf)' AND
TSH_4 = '(-inf-3.15]' AND
on_thyroxine = t: negative (54.0)

referral_source = other AND
T4U_3 = '(1.015-inf)': negative (94.38/1.0)

T3_2 = '(1.85-inf)' AND
TT4_4 = '(131.5-inf)' AND

```

```

on_thyroxine = f AND
referral_source = other AND
T3_3 = '(2.15-inf)': negative (43.09/2.0)

T3_measured = t AND
T3_4 = '(2.45-inf)' AND
T4U_1 = '(0.855-inf)': negative (24.96)

T3_measured = t AND
thyroid_surgery = f AND
T3_2 = '(1.85-inf)' AND
TT4_4 = '(-inf-131.5)': negative (25.09)

T3_measured = t AND
T3_4 = '(-inf-2.45)' AND
thyroid_surgery = f AND
referral_source = other AND
T4U_measured = t AND
age_4 = '(-inf-70.5)' AND
TT4_4 = '(-inf-131.5)' AND
on_thyroxine = f AND
T4U_1 = '(0.855-inf)': negative (39.0)

T3_measured = f: negative (17.78/0.22)

T3_4 = '(-inf-2.45)' AND
thyroid_surgery = f AND
T3_1 = '(1.47-inf)' AND
on_thyroxine = f AND
TT4_4 = '(131.5-inf)' AND
age_1 = '(32.5-inf)' AND
sex = F AND
T4U_2 = '(0.935-inf)': negative (12.57/2.73)

T3_4 = '(-inf-2.45)' AND
thyroid_surgery = f AND
T3_1 = '(1.47-inf)' AND
on_thyroxine = t: negative (11.53)

T3_4 = '(-inf-2.45)' AND
thyroid_surgery = f AND
TT4_measured = t AND
sick = t AND
query_hypothyroid = f AND
T4U_3 = '(-inf-1.015)' AND
T4U_1 = '(-inf-0.855)' AND
FTI_4 = '(-inf-129.5)': sick (14.46/1.46)

T3_4 = '(-inf-2.45)' AND
thyroid_surgery = f AND
TT4_measured = t AND
query_hypothyroid = f AND
referral_source = SVHC AND
FTI_3 = '(113.5-inf)': negative (17.21)

T3_4 = '(-inf-2.45)' AND
thyroid_surgery = f AND
TT4_measured = t AND
query_hypothyroid = f AND
T4U_2 = '(0.935-inf)' AND
age_1 = '(32.5-inf)' AND
query_on_thyroxine = f AND
age_2 = '(47.5-inf)' AND
T3_1 = '(-inf-1.47)' AND
age_3 = '(59.5-inf)' AND
on_thyroxine = f AND
referral_source = SVI AND
TT4_4 = '(-inf-131.5)' AND
FTI_3 = '(-inf-113.5)' AND
TSH_4 = '(3.15-inf)': negative (15.53/1.0)

T3_4 = '(-inf-2.45)' AND
thyroid_surgery = f AND
TT4_measured = t AND
query_hypothyroid = f AND
referral_source = other AND
T4U_measured = t AND
T3_2 = '(-inf-1.85)' AND
sex = M AND
age_4 = '(-inf-70.5)': negative (17.61)

thyroid_surgery = f AND
T3_4 = '(-inf-2.45)' AND
TT4_measured = t AND
TT4_4 = '(131.5-inf)' AND
age_1 = '(32.5-inf)' AND
on_thyroxine = f AND

```

```

T4U_measured = t AND
T3_3 = '(-inf-2.15]' AND
query_on_thyroxine = f AND
query_hypothyroid = f AND
sex = M AND
age_4 = '(-inf-70.5]': negative (8.12)

thyroid_surgery = f AND
T3_4 = '(-inf-2.45]' AND
TT4_measured = t AND
query_hypothyroid = t AND
sick = f AND
age_2 = '(47.5-inf)' AND
TT4_4 = '(-inf-131.5]' AND
TSH_1 = '(0.295-inf)' AND
T4U_4 = '(-inf-1.115]' AND
referral_source = SVI AND
FTI_4 = '(-inf-129.5]' AND
FTI_1 = '(89.5-inf)' AND
TSH_4 = '(-inf-3.15]' AND
FTI_3 = '(-inf-113.5]': sick (7.0/1.0)

thyroid_surgery = f AND
T3_4 = '(-inf-2.45]' AND
TT4_measured = t AND
T4U_1 = '(0.855-inf)' AND
query_hyperthyroid = f AND
T3_2 = '(-inf-1.85]' AND
age_1 = '(32.5-inf)' AND
TT4_1 = '(83.5-inf)' AND
psych = f AND
query_on_thyroxine = f AND
referral_source = SVI AND
T4U_measured = t AND
TSH_4 = '(-inf-3.15]' AND
query_hypothyroid = f AND
FTI_1 = '(89.5-inf)' AND
T4U_4 = '(-inf-1.115]' AND
TSH_1 = '(0.295-inf)' AND
T4U_3 = '(-inf-1.015]' AND
sick = f AND
TSH_2 = '(0.995-inf)' AND
TT4_2 = '(97.5-inf)' AND
FTI_3 = '(113.5-inf)': negative (23.0/4.0)

thyroid_surgery = f AND
T3_4 = '(-inf-2.45]' AND
TT4_measured = t AND
age_1 = '(-inf-32.5]' AND
query_hypothyroid = f AND
T4U_2 = '(-inf-0.935]' AND
TT4_3 = '(-inf-111.5]': negative (17.19/1.0)

thyroid_surgery = t: negative (6.14)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
age_1 = '(-inf-32.5]' AND
T4U_1 = '(0.855-inf)' AND
TSH_3 = '(-inf-1.75]' AND
sex = M: sick (4.38/0.55)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
T4U_2 = '(0.935-inf)' AND
age_1 = '(32.5-inf)' AND
sick = f AND
query_on_thyroxine = f AND
age_2 = '(47.5-inf)' AND
FTI_4 = '(129.5-inf)': negative (9.84/1.46)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
T4U_4 = '(1.115-inf)' AND
FTI_1 = '(-inf-89.5]': negative (8.4/0.15)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
referral_source = other AND
age_1 = '(32.5-inf)' AND
T4U_measured = t AND
TT4_1 = '(83.5-inf)' AND
on_thyroxine = f AND
FTI_1 = '(89.5-inf)' AND
sex = F AND
TSH_4 = '(-inf-3.15]' AND
T3_1 = '(-inf-1.47]': negative (12.18/3.0)

```

```

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
query_on_thyroxine = t AND
TT4_3 = '(111.5-inf)': negative (4.0)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
query_hyperthyroid = t AND
age_2 = '(47.5-inf)' AND
TT4_2 = '(97.5-inf)': sick (3.95/1.0)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
T4U_measured = f AND
referral_source = SVI AND
TT4_1 = '(83.5-inf)': sick (6.58/1.68)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
T4U_measured = t AND
TSH_measured = t AND
T4U_1 = '(-inf-0.855]' AND
age_1 = '(32.5-inf)' AND
T3_1 = '(-inf-1.47)' AND
TT4_4 = '(-inf-131.5]' AND
query_hyperthyroid = f AND
FTI_2 = '(101.5-inf)' AND
TSH_4 = '(3.15-inf)' AND
age_4 = '(-inf-70.5)': sick (17.0/1.0)

TT4_measured = t AND
T3_4 = '(-inf-2.45]' AND
T4U_measured = f: negative (10.61/0.98)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
query_hyperthyroid = t AND
age_2 = '(-inf-47.5)': sick (3.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
TT4_1 = '(83.5-inf)' AND
psych = f AND
FTI_1 = '(-inf-89.5)': negative (6.0/1.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
TT4_1 = '(83.5-inf)' AND
psych = f AND
referral_source = other AND
age_1 = '(32.5-inf)' AND
TT4_2 = '(97.5-inf)' AND
on_thyroxine = f AND
sex = M: negative (5.0/1.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
TT4_1 = '(83.5-inf)' AND
psych = f AND
referral_source = other AND
age_1 = '(32.5-inf)' AND
FTI_4 = '(-inf-129.5)': negative (5.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
query_on_thyroxine = f AND
tumor = f AND
TT4_1 = '(-inf-83.5]' AND
sick = f AND
TSH_4 = '(-inf-3.15]' AND
query_hypothyroid = f AND
on_thyroxine = f AND
TSH_2 = '(0.995-inf)' AND
FTI_2 = '(101.5-inf)': sick (8.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND

```

```

psych = t: negative (7.0/2.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
TSH_measured = t AND
referral_source = SVHC AND
age_2 = '(47.5-inf)': sick (6.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
TSH_measured = t AND
referral_source = other AND
age_1 = '(32.5-inf)' AND
TSH_2 = '(-inf-0.995]' AND
FTI_2 = '(101.5-inf)' AND
T3_1 = '(-inf-1.47]': negative (3.6/1.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
TSH_measured = t AND
referral_source = other AND
age_1 = '(32.5-inf)' AND
FTI_2 = '(-inf-101.5]' AND
TSH_1 = '(0.295-inf)' AND
query_hypothyroid = f AND
T4U_1 = '(0.855-inf)': sick (7.0/1.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
TSH_measured = t AND
referral_source = other AND
FTI_2 = '(-inf-101.5]': negative (8.6/1.0)

T4U_measured = t AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
referral_source = other: sick (8.25/1.25)

referral_source = SVI AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
on_thyroxine = t AND
age_4 = '(70.5-inf)': negative (4.0)

referral_source = SVI AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
T3_1 = '(1.47-inf)' AND
TT4_4 = '(-inf-131.5]': negative (4.0)

referral_source = SVI AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
TSH_1 = '(-inf-0.295]' AND
T4U_2 = '(-inf-0.935]' AND
query_on_thyroxine = f AND
FTI_4 = '(-inf-129.5]' AND
age_2 = '(47.5-inf)' AND
FTI_3 = '(113.5-inf)': sick (11.0/1.0)

referral_source = SVI AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
on_thyroxine = f AND
query_on_thyroxine = f AND
T3_3 = '(-inf-2.15]' AND
TT4_1 = '(83.5-inf)' AND
age_2 = '(-inf-47.5]' AND
FTI_2 = '(101.5-inf)' AND
TSH_3 = '(-inf-1.75]': negative (8.0/2.0)

referral_source = SVI AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
on_thyroxine = f AND
query_on_thyroxine = f AND
T3_3 = '(-inf-2.15]' AND
T3_1 = '(-inf-1.47)' AND

```

```

TT4_1 = '(83.5-inf)' AND
age_2 = '(47.5-inf)' AND
TSH_3 = '(1.75-inf)' AND
sick = f AND
TSH_4 = '(-inf-3.15]' AND
FTI_2 = '(101.5-inf)' AND
query_hypothyroid = f AND
TT4_3 = '(-inf-111.5]' AND
age_4 = '(-inf-70.5]': negative (8.0/1.0)

referral_source = SVI AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
on_thyroxine = f AND
query_on_thyroxine = f AND
T3_3 = '(-inf-2.15]' AND
T3_1 = '(1.47-inf)': negative (4.0/1.0)

referral_source = SVHC: negative (3.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
on_thyroxine = f AND
query_on_thyroxine = f AND
T3_1 = '(-inf-1.47]' AND
TT4_4 = '(131.5-inf)' AND
sex = F: negative (3.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_4 = '(-inf-2.45]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
TT4_4 = '(131.5-inf)': sick (6.0/2.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
on_thyroxine = f AND
query_on_thyroxine = f AND
TT4_1 = '(83.5-inf)' AND
age_2 = '(-inf-47.5]': negative (5.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
age_2 = '(-inf-47.5]' AND
T4U_2 = '(-inf-0.935]': sick (11.0/1.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
referral_source = SVI AND
query_hyperthyroid = f AND
age_2 = '(47.5-inf)' AND
T4U_1 = '(0.855-inf)' AND
TSH_1 = '(0.295-inf)' AND
TT4_3 = '(111.5-inf)' AND
age_3 = '(59.5-inf)': negative (9.0/1.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
T4U_3 = '(1.015-inf)' AND
age_4 = '(-inf-70.5]' AND
age_3 = '(-inf-59.5]': sick (8.0/2.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND

```

```

T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
T4U_1 = '(0.855-inf)' AND
age_3 = '(-inf-59.5]': negative (12.0/2.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
TSH_4 = '(3.15-inf)' AND
FTI_3 = '(-inf-113.5]' AND
TT4_2 = '(-inf-97.5]' AND
T4U_2 = '(-inf-0.935]' AND
T4U_1 = '(-inf-89.5]': negative (22.0/6.0)

T4U_measured = t AND
referral_source = SVI AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
TT4_1 = '(-inf-83.5]' AND
query_hypothyroid = f AND
sex = M AND
TSH_4 = '(-inf-3.15]': sick (20.47/4.47)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
TT4_2 = '(97.5-inf)' AND
age_3 = '(59.5-inf)' AND
query_hypothyroid = f AND
TSH_2 = '(0.995-inf)' AND
T4U_3 = '(-inf-1.015]' AND
sex = F AND
FTI_4 = '(-inf-129.5]': sick (6.56)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
referral_source = SVI AND
query_hyperthyroid = f AND
query_hypothyroid = t AND
TT4_1 = '(-inf-83.5]': sick (4.0)

referral_source = SVI AND
age_1 = '(32.5-inf)' AND
TT4_4 = '(-inf-131.5]' AND
tumor = f AND
query_hyperthyroid = f AND
query_hypothyroid = f AND
FTI_1 = '(-inf-89.5]' AND
age_3 = '(59.5-inf)': negative (7.0/1.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
referral_source = SVI AND
query_hyperthyroid = f AND
query_hypothyroid = f AND
FTI_1 = '(89.5-inf)' AND
FTI_4 = '(129.5-inf)' AND
age_3 = '(59.5-inf)' AND
age_4 = '(70.5-inf)' AND
sick = f AND
TSH_4 = '(-inf-3.15]' AND
TSH_1 = '(0.295-inf)': sick (9.0/3.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
referral_source = SVI AND
query_hyperthyroid = f AND
query_hypothyroid = f AND
FTI_1 = '(89.5-inf)' AND
age_4 = '(70.5-inf)' AND
T4U_3 = '(-inf-1.015]' AND
sick = f AND

```

```

sex = M AND
T4U_1 = '(0.855-inf)': negative (10.0/2.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
query_hypothyroid = f AND
FTI_1 = '(89.5-inf)' AND
FTI_3 = '(113.5-inf)' AND
T4U_1 = '(-inf-0.855]' AND
age_4 = '(-inf-70.5]' AND
TSH_1 = '(0.295-inf)': negative (5.98/2.44)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
query_hypothyroid = f AND
age_3 = '(59.5-inf)' AND
FTI_3 = '(113.5-inf)' AND
age_4 = '(70.5-inf)' AND
sick = f AND
TSH_4 = '(-inf-3.15)': negative (10.0/3.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
query_hypothyroid = f AND
FTI_4 = '(129.5-inf)' AND
age_4 = '(-inf-70.5]': sick (4.0)

T4U_measured = t AND
age_1 = '(32.5-inf)' AND
tumor = f AND
T3_1 = '(-inf-1.47]' AND
query_hyperthyroid = f AND
referral_source = SVI AND
query_hypothyroid = f AND
sick = t AND
age_4 = '(70.5-inf)': negative (4.0)

T4U_measured = t AND
TT4_3 = '(-inf-111.5]' AND
query_hyperthyroid = f AND
sex = M AND
query_hypothyroid = f AND
FTI_2 = '(101.5-inf)': sick (9.26/2.26)

T4U_measured = t AND
TT4_3 = '(111.5-inf)': sick (7.0)

query_hyperthyroid = f AND
age_4 = '(-inf-70.5]' AND
TT4_2 = '(-inf-97.5]' AND
sex = F AND
FTI_2 = '(-inf-101.5]': sick (9.81/2.81)

age_4 = '(-inf-70.5)': negative (14.6/1.0)

sex = F AND
FTI_3 = '(-inf-113.5]' AND
TT4_2 = '(-inf-97.5]' AND
TSH_2 = '(-inf-0.995]': sick (4.25/1.01)

sex = F AND
T4U_3 = '(-inf-1.015]' AND
FTI_3 = '(-inf-113.5]' AND
TT4_1 = '(83.5-inf)' AND
TSH_4 = '(-inf-3.15)': negative (4.01)

T4U_1 = '(0.855-inf)': sick (8.02/0.02)

TT4_1 = '(-inf-83.5)': sick (6.78/2.01)

: negative (3.01)

```

Number of Rules : 75

```

Time taken to build model: 0.33 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3544           93.9555 %
Incorrectly Classified Instances   228            6.0445 %
Kappa statistic                   0.433
Mean absolute error               0.0661
Root mean squared error          0.2153
Relative absolute error          57.3846 %
Root relative squared error     89.762  %
Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall    F-Measure   MCC      ROC Area   PRC Area   Class
      0.973     0.571     0.963     0.973     0.968     0.435     0.880     0.979     negative
      0.429     0.027     0.508     0.429     0.465     0.435     0.880     0.456     sick
Weighted Avg.      0.940     0.538     0.935     0.940     0.937     0.435     0.880     0.947

==== Confusion Matrix ====

      a     b  <-- classified as
3445  96 |   a = negative
 132  99 |   b = sick

```

### 3.2.3 10 bins

#### Cross Validation 5 folds

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B10-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  78
              age_1
              age_2
              age_3
              age_4
              age_5
              age_6
              age_7
              age_8
              age_9
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              TSH_2
              TSH_3
              TSH_4
              TSH_5
              TSH_6
              TSH_7
              TSH_8
              TSH_9
              T3_measured
              T3_1
              T3_2
              T3_3
              T3_4
              T3_5
              T3_6
              T3_7
              T3_8
              T3_9
              TT4_measured
              TT4_1

```

```

TT4_2
TT4_3
TT4_4
TT4_5
TT4_6
TT4_7
TT4_8
TT4_9
T4U_measured
T4U_1
T4U_2
T4U_3
T4U_4
T4U_5
T4U_6
T4U_7
T4U_8
T4U_9
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
FTI_5
FTI_6
FTI_7
FTI_8
FTI_9
TBG_measured
TBG
referral_source
Class
Test mode: 5-fold cross-validation

==== Classifier model (full training set) ====

PART decision list
-----
T3_1 = '(1.15-inf)' AND
TT4_9 = '(-inf-149.5]' AND
TT4_1 = '(72.5-inf)': negative (2774.73)

T3_1 = '(1.15-inf)' AND
referral_source = other AND
FTI_9 = '(-inf-147.5)': negative (247.45/0.6)

T3_1 = '(1.15-inf)' AND
on_thyroxine = f AND
T3_9 = '(-inf-2.85]' AND
TSH_2 = '(0.285-inf)': negative (131.1/4.7)

referral_source = other AND
age_9 = '(-inf-75.5]' AND
T3_measured = f: negative (98.87/0.28)

TT4_9 = '(149.5-inf)' AND
T3_9 = '(2.85-inf)': negative (112.69)

referral_source = other AND
sick = f AND
age_9 = '(-inf-75.5]' AND
T4U_7 = '(1.055-inf)': negative (28.34)

T3_measured = t AND
T3_4 = '(1.85-inf)' AND
on_thyroxine = t: negative (21.66)

T3_measured = t AND
T3_9 = '(-inf-2.85]' AND
TT4_1 = '(72.5-inf)' AND
TT4_9 = '(-inf-149.5]' AND
TSH_measured = t AND
referral_source = SVI AND
FTI_9 = '(-inf-147.5)': sick (121.15/1.0)

T3_measured = f: negative (13.53/1.41)

T3_9 = '(-inf-2.85]' AND
sick = t AND
T3_1 = '(-inf-1.15]': sick (12.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
T4U_measured = t AND
T4U_8 = '(-inf-1.115]' AND
T4U_4 = '(0.935-inf)' AND

```

```

age_1 = '(25.5-inf)' AND
age_8 = '(-inf-70.5]' AND
age_6 = '(59.5-inf)': negative (12.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
T4U_measured = t AND
T4U_8 = '(-inf-1.115]' AND
psych = f AND
TSH_9 = '(6.15-inf)' AND
TT4_1 = '(-inf-72.5]' AND
on_thyroxine = f: negative (34.83/5.55)

T3_9 = '(-inf-2.85]' AND
sick = f AND
T4U_measured = t AND
T4U_8 = '(-inf-1.115]' AND
psych = f AND
TSH_measured = t AND
TSH_8 = '(3.05-inf)' AND
age_5 = '(54.5-inf)' AND
on_thyroxine = f: sick (12.0/2.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
T4U_measured = f: negative (10.83/0.85)

T3_9 = '(-inf-2.85]' AND
T4U_8 = '(1.115-inf)': negative (7.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
psych = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
on_thyroxine = t AND
age_5 = '(54.5-inf)' AND
TSH_3 = '(-inf-0.655)': negative (6.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
psych = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
T4U_4 = '(-inf-0.935]' AND
T3_6 = '(-inf-2.05]' AND
T3_4 = '(-inf-1.85]' AND
T3_1 = '(-inf-1.15]' AND
age_1 = '(25.5-inf)' AND
TSH_1 = '(0.085-inf)' AND
TT4_8 = '(-inf-130.5]' AND
age_7 = '(64.5-inf)' AND
TSH_3 = '(0.655-inf)' AND
FTI_1 = '(78.5-inf)': sick (20.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
psych = f AND
T4U_1 = '(0.785-inf)' AND
TSH_8 = '(-inf-3.05]' AND
query_hyperthyroid = f AND
T4U_3 = '(-inf-0.895]' AND
TT4_9 = '(-inf-149.5]' AND
FTI_1 = '(78.5-inf)': negative (10.0/3.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
psych = f AND
T4U_1 = '(0.785-inf)' AND
T4U_3 = '(-inf-0.895)': sick (20.72)

T3_7 = '(2.25-inf)': negative (6.39)

psych = f AND
sick = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
T4U_7 = '(-inf-1.055]' AND
T3_2 = '(-inf-1.47]' AND
TT4_7 = '(-inf-119.5]' AND
FTI_7 = '(120.5-inf)' AND
TSH_2 = '(-inf-0.285)': negative (4.0)

psych = f AND
sick = f AND
TSH_measured = t AND
query_hyperthyroid = f AND

```

```

TSH_3 = '(0.655-inf)' AND
query_hypothyroid = f AND
on_thyroxine = f AND
TT4_4 = '(-inf-96.5]' AND
age_1 = '(25.5-inf)' AND
age_4 = '(-inf-47.5]': negative (5.0/1.0)

psych = f AND
sick = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
TSH_3 = '(0.655-inf)' AND
FTI_6 = '(-inf-113.5]' AND
query_hypothyroid = f AND
on_thyroxine = f AND
age_4 = '(47.5-inf)' AND
T4U_1 = '(-inf-0.785]' AND
FTI_1 = '(-inf-78.5]': negative (5.0/1.0)

sick = f AND
psych = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
TSH_1 = '(-inf-0.085)' AND
age_9 = '(-inf-75.5]' AND
referral_source = SVI AND
T4U_1 = '(0.785-inf)': negative (4.0)

sick = f AND
T3_5 = '(-inf-1.95]' AND
psych = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
age_6 = '(-inf-59.5]' AND
FTI_2 = '(-inf-89.5]': sick (7.0)

T3_5 = '(-inf-1.95]' AND
psych = f AND
T4U_6 = '(-inf-1.015]' AND
query_on_thyroxine = f AND
sick = f AND
T3_1 = '(1.15-inf)': sick (6.0/1.0)

T3_5 = '(-inf-1.95]' AND
psych = f AND
T3_1 = '(1.15-inf)': negative (4.26)

T3_1 = '(-inf-1.15)' AND
psych = f AND
age_9 = '(75.5-inf)' AND
T4U_4 = '(0.935-inf)': negative (3.0/1.0)

TT4_8 = '(-inf-130.5]' AND
age_9 = '(-inf-75.5]' AND
psych = f AND
referral_source = other: negative (9.72/1.0)

T4U_1 = '(-inf-0.785]' AND
TSH_2 = '(0.285-inf)' AND
FTI_3 = '(95.5-inf)': sick (7.5/3.0)

T4U_1 = '(0.785-inf)': sick (6.0)

TSH_2 = '(-inf-0.285]': sick (5.13/1.0)

: negative (4.1/1.1)

Number of Rules : 33

Time taken to build model: 0.26 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3702          98.1442 %
Incorrectly Classified Instances    70           1.8558 %
Kappa statistic                   0.8359
Mean absolute error                0.0228
Root mean squared error            0.1267
Relative absolute error             19.8065 %
Root relative squared error       52.8442 %
Total Number of Instances         3772

==== Detailed Accuracy By Class ====

```

TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
---------	---------	-----------	--------	-----------	-----	----------	----------	-------

```

0.991    0.169    0.989    0.991    0.990    0.836    0.977    0.998    negative
0.831    0.009    0.861    0.831    0.846    0.836    0.977    0.868    sick
Weighted Avg. 0.981    0.159    0.981    0.981    0.981    0.836    0.977    0.990

==== Confusion Matrix ====
a      b    <-- classified as
3510   31 |     a = negative
 39  192 |     b = sick

```

## Cross Validation 10 folds

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B10-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  78
              age_1
              age_2
              age_3
              age_4
              age_5
              age_6
              age_7
              age_8
              age_9
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              t131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              TSH_2
              TSH_3
              TSH_4
              TSH_5
              TSH_6
              TSH_7
              TSH_8
              TSH_9
              T3_measured
              T3_1
              T3_2
              T3_3
              T3_4
              T3_5
              T3_6
              T3_7
              T3_8
              T3_9
              TT4_measured
              TT4_1
              TT4_2
              TT4_3
              TT4_4
              TT4_5
              TT4_6
              TT4_7
              TT4_8
              TT4_9
              T4U_measured
              T4U_1
              T4U_2
              T4U_3
              T4U_4
              T4U_5
              T4U_6
              T4U_7
              T4U_8
              T4U_9
              FTI_measured
              FTI_1

```

```

FTI_2
FTI_3
FTI_4
FTI_5
FTI_6
FTI_7
FTI_8
FTI_9
TBG_measured
TBG
referral_source
Class
Test mode: 10-fold cross-validation

==== Classifier model (full training set) ====

PART decision list
-----

T3_1 = '(1.15-inf)' AND
TT4_9 = '(-inf-149.5]' AND
TT4_1 = '(72.5-inf)': negative (2774.73)

T3_1 = '(1.15-inf)' AND
referral_source = other AND
FTI_9 = '(-inf-147.5)': negative (247.45/0.6)

T3_1 = '(1.15-inf)' AND
on_thyroxine = f AND
T3_9 = '(-inf-2.85]' AND
TSH_2 = '(0.285-inf)': negative (131.1/4.7)

referral_source = other AND
age_9 = '(-inf-75.5]' AND
T3_measured = f: negative (98.87/0.28)

TT4_9 = '(149.5-inf)' AND
T3_9 = '(2.85-inf)': negative (112.69)

referral_source = other AND
sick = f AND
age_9 = '(-inf-75.5]' AND
T4U_7 = '(1.055-inf)': negative (28.34)

T3_measured = t AND
T3_4 = '(1.85-inf)' AND
on_thyroxine = t: negative (21.66)

T3_measured = t AND
T3_9 = '(-inf-2.85]' AND
TT4_1 = '(72.5-inf)' AND
TT4_9 = '(-inf-149.5]' AND
TSH_measured = t AND
referral_source = SVI AND
FTI_9 = '(-inf-147.5)': sick (121.15/1.0)

T3_measured = f: negative (13.53/1.41)

T3_9 = '(-inf-2.85]' AND
sick = t AND
T3_1 = '(-inf-1.15)': sick (12.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
T4U_measured = t AND
T4U_8 = '(-inf-1.115]' AND
T4U_4 = '(0.935-inf)' AND
age_1 = '(25.5-inf)' AND
age_8 = '(-inf-70.5]' AND
age_6 = '(59.5-inf)': negative (12.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
T4U_measured = t AND
T4U_8 = '(-inf-1.115]' AND
psych = f AND
TSH_9 = '(6.15-inf)' AND
TT4_1 = '(-inf-72.5]' AND
on_thyroxine = f: negative (34.83/5.55)

T3_9 = '(-inf-2.85]' AND
sick = f AND
T4U_measured = t AND
T4U_8 = '(-inf-1.115]' AND
psych = f AND
TSH_measured = t AND
TSH_8 = '(3.05-inf)' AND

```

```

age_5 = '(54.5-inf)' AND
on_thyroxine = f: sick (12.0/2.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
T4U_measured = f: negative (10.83/0.85)

T3_9 = '(-inf-2.85]' AND
T4U_8 = '(1.115-inf)': negative (7.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
psych = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
on_thyroxine = t AND
age_5 = '(54.5-inf)' AND
TSH_3 = '(-inf-0.655)': negative (6.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
psych = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
T4U_4 = '(-inf-0.935]' AND
T3_6 = '(-inf-2.05]' AND
T3_4 = '(-inf-1.85]' AND
T3_1 = '(-inf-1.15]' AND
age_1 = '(25.5-inf)' AND
TSH_1 = '(0.085-inf)' AND
TT4_8 = '(-inf-130.5]' AND
age_7 = '(64.5-inf)' AND
TSH_3 = '(0.655-inf)' AND
FTI_1 = '(78.5-inf)': sick (20.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
psych = f AND
T4U_1 = '(0.785-inf)' AND
TSH_8 = '(-inf-3.05]' AND
query_hyperthyroid = f AND
T4U_3 = '(-inf-0.895]' AND
TT4_9 = '(-inf-149.5]' AND
FTI_1 = '(78.5-inf)': negative (10.0/3.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
psych = f AND
T4U_1 = '(0.785-inf)' AND
T4U_3 = '(-inf-0.895)': sick (20.72)

T3_7 = '(2.25-inf)': negative (6.39)

psych = f AND
sick = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
T4U_7 = '(-inf-1.055]' AND
T3_2 = '(-inf-1.47]' AND
TT4_7 = '(-inf-119.5]' AND
FTI_7 = '(120.5-inf)' AND
TSH_2 = '(-inf-0.285)': negative (4.0)

psych = f AND
sick = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
TSH_3 = '(0.655-inf)' AND
query_hypothyroid = f AND
on_thyroxine = f AND
age_1 = '(25.5-inf)' AND
age_4 = '(-inf-47.5)': negative (5.0/1.0)

psych = f AND
sick = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
TSH_3 = '(0.655-inf)' AND
FTI_6 = '(-inf-113.5]' AND
query_hypothyroid = f AND
on_thyroxine = f AND
age_4 = '(47.5-inf)' AND
T4U_1 = '(-inf-0.785]' AND
FTI_1 = '(-inf-78.5)': negative (5.0/1.0)

sick = f AND

```

```

psych = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
TSH_1 = '(-inf-0.085]' AND
age_9 = '(-inf-75.5]' AND
referral_source = SVI AND
T4U_1 = '(0.785-inf)': negative (4.0)

sick = f AND
T3_5 = '(-inf-1.95]' AND
psych = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
age_6 = '(-inf-59.5]' AND
FTI_2 = '(-inf-89.5]': sick (7.0)

T3_5 = '(-inf-1.95]' AND
psych = f AND
T4U_6 = '(-inf-1.015]' AND
query_on_thyroxine = f AND
sick = f AND
T3_1 = '(1.15-inf)': sick (6.0/1.0)

T3_5 = '(-inf-1.95]' AND
psych = f AND
T3_1 = '(1.15-inf)': negative (4.26)

T3_1 = '(-inf-1.15]' AND
psych = f AND
age_9 = '(75.5-inf)' AND
T4U_4 = '(0.935-inf)': negative (3.0/1.0)

TT4_8 = '(-inf-130.5]' AND
age_9 = '(-inf-75.5]' AND
psych = f AND
referral_source = other: negative (9.72/1.0)

T4U_1 = '(-inf-0.785]' AND
TSH_2 = '(0.285-inf)' AND
FTI_3 = '(95.5-inf)': sick (7.5/3.0)

T4U_1 = '(0.785-inf)': sick (6.0)

TSH_2 = '(-inf-0.285)': sick (5.13/1.0)

: negative (4.1/1.1)

Number of Rules : 33

Time taken to build model: 0.2 seconds

==== Stratified cross-validation ====
==== Summary ====

  Correctly Classified Instances      3687      97.7466 %
  Incorrectly Classified Instances     85       2.2534 %
  Kappa statistic                      0.8052
  Mean absolute error                  0.0252
  Root mean squared error              0.1343
  Relative absolute error              21.8808 %
  Root relative squared error          56.0012 %
  Total Number of Instances            3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.988     0.177     0.988     0.988     0.988     0.805     0.966     0.996   negative
      0.823     0.012     0.812     0.823     0.817     0.805     0.966     0.830     sick
  Weighted Avg.      0.977     0.167     0.978     0.977     0.978     0.805     0.966     0.986

==== Confusion Matrix ====

      a      b  <-- classified as
  3497    44 |      a = negative
      41   190 |      b = sick

```

## Cross Validation 15 folds

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize=D-F=B10-M=1.0-Rfirst-last-precision6
Instances:   3772

```

```
Attributes:    78
              age_1
              age_2
              age_3
              age_4
              age_5
              age_6
              age_7
              age_8
              age_9
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              t131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              TSH_2
              TSH_3
              TSH_4
              TSH_5
              TSH_6
              TSH_7
              TSH_8
              TSH_9
              T3_measured
              T3_1
              T3_2
              T3_3
              T3_4
              T3_5
              T3_6
              T3_7
              T3_8
              T3_9
              TT4_measured
              TT4_1
              TT4_2
              TT4_3
              TT4_4
              TT4_5
              TT4_6
              TT4_7
              TT4_8
              TT4_9
              T4U_measured
              T4U_1
              T4U_2
              T4U_3
              T4U_4
              T4U_5
              T4U_6
              T4U_7
              T4U_8
              T4U_9
              FTI_measured
              FTI_1
              FTI_2
              FTI_3
              FTI_4
              FTI_5
              FTI_6
              FTI_7
              FTI_8
              FTI_9
              TBG_measured
              TBG
              referral_source
              Class

Test mode: 15-fold cross-validation

==== Classifier model (full training set) ====

PART decision list
-----
T3_1 = '(1.15-inf)' AND
```

```

TT4_9 = '(-inf-149.5]' AND
TT4_1 = '(72.5-inf)': negative (2774.73)

T3_1 = '(1.15-inf)' AND
referral_source = other AND
FTI_9 = '(-inf-147.5]': negative (247.45/0.6)

T3_1 = '(1.15-inf)' AND
on_thyroxine = f AND
T3_9 = '(-inf-2.85]' AND
TSH_2 = '(0.285-inf)': negative (131.1/4.7)

referral_source = other AND
age_9 = '(-inf-75.5]' AND
T3_measured = f: negative (98.87/0.28)

TT4_9 = '(149.5-inf)' AND
T3_9 = '(2.85-inf)': negative (112.69)

referral_source = other AND
sick = f AND
age_9 = '(-inf-75.5]' AND
T4U_7 = '(1.055-inf)': negative (28.34)

T3_measured = t AND
T3_4 = '(1.85-inf)' AND
on_thyroxine = t: negative (21.66)

T3_measured = t AND
T3_9 = '(-inf-2.85]' AND
TT4_1 = '(72.5-inf)' AND
TT4_9 = '(-inf-149.5]' AND
TSH_measured = t AND
referral_source = SVI AND
FTI_9 = '(-inf-147.5]': sick (121.15/1.0)

T3_measured = f: negative (13.53/1.41)

T3_9 = '(-inf-2.85]' AND
sick = t AND
T3_1 = '(-inf-1.15]': sick (12.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
T4U_measured = t AND
T4U_8 = '(-inf-1.115]' AND
T4U_4 = '(0.935-inf)' AND
age_1 = '(25.5-inf)' AND
age_8 = '(-inf-70.5]' AND
age_6 = '(59.5-inf)': negative (12.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
T4U_measured = t AND
T4U_8 = '(-inf-1.115]' AND
psych = f AND
TSH_9 = '(6.15-inf)' AND
TT4_1 = '(-inf-72.5]' AND
on_thyroxine = f: negative (34.83/5.55)

T3_9 = '(-inf-2.85]' AND
sick = f AND
T4U_measured = t AND
T4U_8 = '(-inf-1.115]' AND
psych = f AND
TSH_measured = t AND
TSH_8 = '(3.05-inf)' AND
age_5 = '(54.5-inf)' AND
on_thyroxine = f: sick (12.0/2.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
T4U_measured = f: negative (10.83/0.85)

T3_9 = '(-inf-2.85]' AND
T4U_8 = '(1.115-inf)': negative (7.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
psych = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
on_thyroxine = t AND
age_5 = '(54.5-inf)' AND
TSH_3 = '(-inf-0.655]': negative (6.0)

T3_9 = '(-inf-2.85]' AND

```

```

sick = f AND
psych = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
T4U_4 = '(-inf-0.935]' AND
T3_6 = '(-inf-2.05]' AND
T3_4 = '(-inf-1.85]' AND
T3_1 = '(-inf-1.15]' AND
age_1 = '(25.5-inf)' AND
TSH_1 = '(0.085-inf)' AND
TT4_8 = '(-inf-130.5]' AND
age_7 = '(64.5-inf)' AND
TSH_3 = '(0.655-inf)' AND
FTI_1 = '(78.5-inf)': sick (20.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
psych = f AND
T4U_1 = '(0.785-inf)' AND
TSH_8 = '(-inf-3.05]' AND
query_hyperthyroid = f AND
T4U_3 = '(-inf-0.895]' AND
TT4_9 = '(-inf-149.5]' AND
FTI_1 = '(78.5-inf)': negative (10.0/3.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
psych = f AND
T4U_1 = '(0.785-inf)' AND
T4U_3 = '(-inf-0.895)': sick (20.72)

T3_7 = '(2.25-inf)': negative (6.39)

psych = f AND
sick = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
T4U_7 = '(-inf-1.055]' AND
T3_2 = '(-inf-1.47]' AND
TT4_7 = '(-inf-119.5]' AND
FTI_7 = '(120.5-inf)' AND
TSH_2 = '(-inf-0.285)': negative (4.0)

psych = f AND
sick = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
TSH_3 = '(0.655-inf)' AND
query_hypothyroid = f AND
on_thyroxine = f AND
TT4_4 = '(-inf-96.5]' AND
age_1 = '(25.5-inf)' AND
age_4 = '(-inf-47.5)': negative (5.0/1.0)

psych = f AND
sick = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
TSH_3 = '(0.655-inf)' AND
FTI_6 = '(-inf-113.5]' AND
query_hypothyroid = f AND
on_thyroxine = f AND
age_4 = '(47.5-inf)' AND
T4U_1 = '(-inf-0.785]' AND
FTI_1 = '(-inf-78.5)': negative (5.0/1.0)

sick = f AND
psych = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
TSH_1 = '(-inf-0.085)' AND
age_9 = '(-inf-75.5]' AND
referral_source = SVI AND
T4U_1 = '(0.785-inf)': negative (4.0)

sick = f AND
T3_5 = '(-inf-1.95]' AND
psych = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
age_6 = '(-inf-59.5]' AND
FTI_2 = '(-inf-89.5)': sick (7.0)

T3_5 = '(-inf-1.95]' AND
psych = f AND
T4U_6 = '(-inf-1.015]' AND
query_on_thyroxine = f AND

```

```

sick = f AND
T3_1 = '(1.15-inf)': sick (6.0/1.0)

T3_5 = '(-inf-1.95]' AND
psych = f AND
T3_1 = '(1.15-inf)': negative (4.26)

T3_1 = '(-inf-1.15]' AND
psych = f AND
age_9 = '(75.5-inf)' AND
T4U_4 = '(0.935-inf)': negative (3.0/1.0)

TT4_8 = '(-inf-130.5]' AND
age_9 = '(-inf-75.5]' AND
psych = f AND
referral_source = other: negative (9.72/1.0)

T4U_1 = '(-inf-0.785]' AND
TSH_2 = '(0.285-inf)' AND
FTI_3 = '(95.5-inf)': sick (7.5/3.0)

T4U_1 = '(0.785-inf)': sick (6.0)

TSH_2 = '(-inf-0.285]': sick (5.13/1.0)

: negative (4.1/1.1)

Number of Rules : 33

Time taken to build model: 0.21 seconds

==== Stratified cross-validation ====
==== Summary ====

    Correctly Classified Instances      3693          97.9056 %
    Incorrectly Classified Instances   79           2.0944 %
    Kappa statistic                   0.8175
    Mean absolute error              0.0248
    Root mean squared error          0.1364
    Relative absolute error          21.5086 %
    Root relative squared error     56.8865 %
    Total Number of Instances        3772

==== Detailed Accuracy By Class ====

    TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
    0.989     0.173     0.989     0.989     0.989     0.817     0.963     0.996     negative
    0.827     0.011     0.830     0.827     0.829     0.817     0.963     0.789     sick
    Weighted Avg.   0.979     0.163     0.979     0.979     0.979     0.817     0.963     0.983

==== Confusion Matrix ====

    a     b  <-- classified as
3502  39 |    a = negative
    40  191 |    b = sick

```

## Percentage Split 50%

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B10-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  78
            age_1
            age_2
            age_3
            age_4
            age_5
            age_6
            age_7
            age_8
            age_9
            sex
            on_thyroxine
            query_on_thyroxine
            on_antithyroid_medication
            sick
            pregnant
            thyroid_surgery
            t131_treatment
            query_hypothyroid
            query_hyperthyroid

```

```

lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH_1
TSH_2
TSH_3
TSH_4
TSH_5
TSH_6
TSH_7
TSH_8
TSH_9
T3_measured
T3_1
T3_2
T3_3
T3_4
T3_5
T3_6
T3_7
T3_8
T3_9
TT4_measured
TT4_1
TT4_2
TT4_3
TT4_4
TT4_5
TT4_6
TT4_7
TT4_8
TT4_9
T4U_measured
T4U_1
T4U_2
T4U_3
T4U_4
T4U_5
T4U_6
T4U_7
T4U_8
T4U_9
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
FTI_5
FTI_6
FTI_7
FTI_8
FTI_9
TBG_measured
TBG
referral_source
Class
Test mode: split 50.0% train, remainder test

==== Classifier model (full training set) ===

PART decision list
-----
T3_1 = '(1.15-inf)' AND
TT4_9 = '(-inf-149.5]' AND
TT4_1 = '(72.5-inf)': negative (2774.73)

T3_1 = '(1.15-inf)' AND
referral_source = other AND
FTI_9 = '(-inf-147.5]': negative (247.45/0.6)

T3_1 = '(1.15-inf)' AND
on_thyroxine = f AND
T3_9 = '(-inf-2.85]' AND
TSH_2 = '(0.285-inf)': negative (131.1/4.7)

referral_source = other AND
age_9 = '(-inf-75.5]' AND
T3_measured = f: negative (98.87/0.28)

TT4_9 = '(149.5-inf)' AND
T3_9 = '(2.85-inf)': negative (112.69)

referral_source = other AND

```

```

sick = f AND
age_9 = '(-inf-75.5]' AND
T4U_7 = '(1.055-inf)': negative (28.34)

T3_measured = t AND
T3_4 = '(1.85-inf)' AND
on_thyroxine = t: negative (21.66)

T3_measured = t AND
T3_9 = '(-inf-2.85]' AND
TT4_1 = '(72.5-inf)' AND
TT4_9 = '(-inf-149.5]' AND
TSH_measured = t AND
referral_source = SVI AND
FTI_9 = '(-inf-147.5)': sick (121.15/1.0)

T3_measured = f: negative (13.53/1.41)

T3_9 = '(-inf-2.85]' AND
sick = t AND
T3_1 = '(-inf-1.15)': sick (12.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
T4U_measured = t AND
T4U_8 = '(-inf-1.115]' AND
T4U_4 = '(0.935-inf)' AND
age_1 = '(25.5-inf)' AND
age_8 = '(-inf-70.5]' AND
age_6 = '(59.5-inf)': negative (12.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
T4U_measured = t AND
T4U_8 = '(-inf-1.115]' AND
psych = f AND
TSH_9 = '(6.15-inf)' AND
TT4_1 = '(-inf-72.5]' AND
on_thyroxine = f: negative (34.83/5.55)

T3_9 = '(-inf-2.85]' AND
sick = f AND
T4U_measured = t AND
T4U_8 = '(-inf-1.115]' AND
psych = f AND
TSH_measured = t AND
TSH_8 = '(3.05-inf)' AND
age_5 = '(54.5-inf)' AND
on_thyroxine = f: sick (12.0/2.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
T4U_measured = f: negative (10.83/0.85)

T3_9 = '(-inf-2.85]' AND
T4U_8 = '(1.115-inf)': negative (7.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
psych = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
on_thyroxine = t AND
age_5 = '(54.5-inf)' AND
TSH_3 = '(-inf-0.655)': negative (6.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
psych = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
T4U_4 = '(-inf-0.935]' AND
T3_6 = '(-inf-2.05]' AND
T3_4 = '(-inf-1.85]' AND
T3_1 = '(-inf-1.15]' AND
age_1 = '(25.5-inf)' AND
TSH_1 = '(0.085-inf)' AND
TT4_8 = '(-inf-130.5]' AND
age_7 = '(64.5-inf)' AND
TSH_3 = '(0.655-inf)' AND
FTI_1 = '(78.5-inf)': sick (20.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
psych = f AND
T4U_1 = '(0.785-inf)' AND
TSH_8 = '(-inf-3.05]' AND

```

```

query_hyperthyroid = f AND
T4U_3 = '(-inf-0.895]' AND
TT4_9 = '(-inf-149.5]' AND
FTI_1 = '(78.5-inf)': negative (10.0/3.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
psych = f AND
T4U_1 = '(0.785-inf)' AND
T4U_3 = '(-inf-0.895)': sick (20.72)

T3_7 = '(2.25-inf)': negative (6.39)

psych = f AND
sick = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
T4U_7 = '(-inf-1.055]' AND
T3_2 = '(-inf-1.47]' AND
TT4_7 = '(-inf-119.5]' AND
FTI_7 = '(120.5-inf)' AND
TSH_2 = '(-inf-0.285)': negative (4.0)

psych = f AND
sick = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
TSH_3 = '(0.655-inf)' AND
query_hypothyroid = f AND
on_thyroxine = f AND
TT4_4 = '(-inf-96.5]' AND
age_1 = '(25.5-inf)' AND
age_4 = '(-inf-47.5)': negative (5.0/1.0)

psych = f AND
sick = f AND
TSH_measured = t AND
query_hypothyroid = f AND
TSH_3 = '(0.655-inf)' AND
FTI_6 = '(-inf-113.5]' AND
query_hypothyroid = f AND
on_thyroxine = f AND
age_4 = '(47.5-inf)' AND
T4U_1 = '(-inf-0.785)' AND
FTI_1 = '(-inf-78.5)': negative (5.0/1.0)

sick = f AND
psych = f AND
TSH_measured = t AND
query_hypothyroid = f AND
TSH_1 = '(-inf-0.085)' AND
age_9 = '(-inf-75.5]' AND
referral_source = SVI AND
T4U_1 = '(0.785-inf)': negative (4.0)

sick = f AND
T3_5 = '(-inf-1.95]' AND
psych = f AND
TSH_measured = t AND
query_hypothyroid = f AND
age_6 = '(-inf-59.5]' AND
FTI_2 = '(-inf-89.5)': sick (7.0)

T3_5 = '(-inf-1.95)' AND
psych = f AND
T4U_6 = '(-inf-1.015)' AND
query_on_thyroxine = f AND
sick = f AND
T3_1 = '(1.15-inf)': sick (6.0/1.0)

T3_5 = '(-inf-1.95)' AND
psych = f AND
T3_1 = '(1.15-inf)': negative (4.26)

T3_1 = '(-inf-1.15)' AND
psych = f AND
age_9 = '(75.5-inf)' AND
T4U_4 = '(0.935-inf)': negative (3.0/1.0)

TT4_8 = '(-inf-130.5)' AND
age_9 = '(-inf-75.5]' AND
psych = f AND
referral_source = other: negative (9.72/1.0)

T4U_1 = '(-inf-0.785)' AND
TSH_2 = '(0.285-inf)' AND
FTI_3 = '(95.5-inf)': sick (7.5/3.0)

```

```

T4U_1 = '(0.785-inf)': sick (6.0)
TSH_2 = '(-inf-0.285)': sick (5.13/1.0)
: negative (4.1/1.1)

Number of Rules : 33

Time taken to build model: 0.22 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0.01 seconds

==== Summary ===

Correctly Classified Instances      1845          97.8261 %
Incorrectly Classified Instances   41           2.1739 %
Kappa statistic                   0.8094
Mean absolute error               0.0252
Root mean squared error          0.1405
Relative absolute error          21.8363 %
Root relative squared error     58.2538 %
Total Number of Instances        1886

==== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC    ROC Area  PRC Area  Class
0.990     0.197    0.987    0.990    0.988    0.810    0.938    0.992    negative
0.803     0.010    0.839    0.803    0.821    0.810    0.939    0.818    sick
Weighted Avg.  0.978    0.185    0.978    0.978    0.978    0.810    0.938    0.981

==== Confusion Matrix ===

     a     b  <-- classified as
1751  18 |     a = negative
     23  94 |     b = sick

```

## Percentage Split 60%

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B10-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  78
              age_1
              age_2
              age_3
              age_4
              age_5
              age_6
              age_7
              age_8
              age_9
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              TSH_2
              TSH_3
              TSH_4
              TSH_5
              TSH_6
              TSH_7
              TSH_8
              TSH_9
              T3_measured
              T3_1

```

```

T3_2
T3_3
T3_4
T3_5
T3_6
T3_7
T3_8
T3_9
TT4_measured
TT4_1
TT4_2
TT4_3
TT4_4
TT4_5
TT4_6
TT4_7
TT4_8
TT4_9
T4U_measured
T4U_1
T4U_2
T4U_3
T4U_4
T4U_5
T4U_6
T4U_7
T4U_8
T4U_9
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
FTI_5
FTI_6
FTI_7
FTI_8
FTI_9
TBG_measured
TBG
referral_source
Class
Test mode: split 60.0% train, remainder test

==== Classifier model (full training set) ====

PART decision list
-----
T3_1 = '(1.15-inf)' AND
TT4_9 = '(-inf-149.5]' AND
TT4_1 = '(72.5-inf)': negative (2774.73)

T3_1 = '(1.15-inf)' AND
referral_source = other AND
FTI_9 = '(-inf-147.5]': negative (247.45/0.6)

T3_1 = '(1.15-inf)' AND
on_thyroxine = f AND
T3_9 = '(-inf-2.85]' AND
TSH_2 = '(0.285-inf)': negative (131.1/4.7)

referral_source = other AND
age_9 = '(-inf-75.5]' AND
T3_measured = f: negative (98.87/0.28)

TT4_9 = '(149.5-inf)' AND
T3_9 = '(2.85-inf)': negative (112.69)

referral_source = other AND
sick = f AND
age_9 = '(-inf-75.5]' AND
T4U_7 = '(1.055-inf)': negative (28.34)

T3_measured = t AND
T3_4 = '(1.85-inf)' AND
on_thyroxine = t: negative (21.66)

T3_measured = t AND
T3_9 = '(-inf-2.85]' AND
TT4_1 = '(72.5-inf)' AND
TT4_9 = '(-inf-149.5]' AND
TSH_measured = t AND
referral_source = SVI AND
FTI_9 = '(-inf-147.5]': sick (121.15/1.0)

T3_measured = f: negative (13.53/1.41)

```

```

T3_9 = '(-inf-2.85]' AND
sick = t AND
T3_1 = '(-inf-1.15]': sick (12.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
T4U_measured = t AND
T4U_8 = '(-inf-1.115]' AND
T4U_4 = '(0.935-inf)' AND
age_1 = '(25.5-inf)' AND
age_8 = '(-inf-70.5]' AND
age_6 = '(59.5-inf)': negative (12.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
T4U_measured = t AND
T4U_8 = '(-inf-1.115]' AND
psych = f AND
TSH_9 = '(6.15-inf)' AND
TT4_1 = '(-inf-72.5]' AND
on_thyroxine = f: negative (34.83/5.55)

T3_9 = '(-inf-2.85]' AND
sick = f AND
T4U_measured = t AND
T4U_8 = '(-inf-1.115]' AND
psych = f AND
TSH_measured = t AND
TSH_8 = '(3.05-inf)' AND
age_5 = '(54.5-inf)' AND
on_thyroxine = f: sick (12.0/2.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
T4U_measured = f: negative (10.83/0.85)

T3_9 = '(-inf-2.85]' AND
T4U_8 = '(1.115-inf)': negative (7.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
psych = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
on_thyroxine = t AND
age_5 = '(54.5-inf)' AND
TSH_3 = '(-inf-0.655]': negative (6.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
psych = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
T4U_4 = '(-inf-0.935]' AND
T3_6 = '(-inf-2.05]' AND
T3_4 = '(-inf-1.85]' AND
T3_1 = '(-inf-1.15]' AND
age_1 = '(25.5-inf)' AND
TSH_1 = '(0.085-inf)' AND
TT4_8 = '(-inf-130.5]' AND
age_7 = '(64.5-inf)' AND
TSH_3 = '(0.655-inf)' AND
FTI_1 = '(78.5-inf)': sick (20.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
psych = f AND
T4U_1 = '(0.785-inf)' AND
TSH_8 = '(-inf-3.05]' AND
query_hyperthyroid = f AND
T4U_3 = '(-inf-0.895]' AND
TT4_9 = '(-inf-149.5]' AND
FTI_1 = '(78.5-inf)': negative (10.0/3.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
psych = f AND
T4U_1 = '(0.785-inf)' AND
T4U_3 = '(-inf-0.895]': sick (20.72)

T3_7 = '(2.25-inf)': negative (6.39)

psych = f AND
sick = f AND
TSH_measured = t AND
query_hyperthyroid = f AND

```

```

T4U_7 = '(-inf-1.055]' AND
T3_2 = '(-inf-1.47)' AND
TT4_7 = '(-inf-119.5]' AND
FTI_7 = '(120.5-inf)' AND
TSH_2 = '(-inf-0.285]': negative (4.0)

psych = f AND
sick = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
TSH_3 = '(0.655-inf)' AND
query_hypothyroid = f AND
on_thyroxine = f AND
TT4_4 = '(-inf-96.5]' AND
age_1 = '(25.5-inf)' AND
age_4 = '(-inf-47.5]': negative (5.0/1.0)

psych = f AND
sick = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
TSH_3 = '(0.655-inf)' AND
FTI_6 = '(-inf-113.5]' AND
query_hypothyroid = f AND
on_thyroxine = f AND
age_4 = '(47.5-inf)' AND
T4U_1 = '(-inf-0.785]' AND
FTI_1 = '(-inf-78.5]': negative (5.0/1.0)

sick = f AND
psych = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
TSH_1 = '(-inf-0.085]' AND
age_9 = '(-inf-75.5]' AND
referral_source = SVI AND
T4U_1 = '(0.785-inf)': negative (4.0)

sick = f AND
T3_5 = '(-inf-1.95]' AND
psych = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
age_6 = '(-inf-59.5]' AND
FTI_2 = '(-inf-89.5]': sick (7.0)

T3_5 = '(-inf-1.95]' AND
psych = f AND
T4U_6 = '(-inf-1.015]' AND
query_on_thyroxine = f AND
sick = f AND
T3_1 = '(1.15-inf)': sick (6.0/1.0)

T3_5 = '(-inf-1.95]' AND
psych = f AND
T3_1 = '(1.15-inf)': negative (4.26)

T3_1 = '(-inf-1.15]' AND
psych = f AND
age_9 = '(75.5-inf)' AND
T4U_4 = '(0.935-inf)': negative (3.0/1.0)

TT4_8 = '(-inf-130.5]' AND
age_9 = '(-inf-75.5]' AND
psych = f AND
referral_source = other: negative (9.72/1.0)

T4U_1 = '(-inf-0.785]' AND
TSH_2 = '(0.285-inf)' AND
FTI_3 = '(95.5-inf)': sick (7.5/3.0)

T4U_1 = '(0.785-inf)': sick (6.0)

TSH_2 = '(-inf-0.285]': sick (5.13/1.0)

: negative (4.1/1.1)

Number of Rules : 33

```

```

Time taken to build model: 0.22 seconds
==== Evaluation on test split ====
Time taken to test model on test split: 0 seconds
==== Summary ====

```

```

Correctly Classified Instances      1468          97.283 %
Incorrectly Classified Instances   41           2.717 %
Kappa statistic                   0.7513
Mean absolute error               0.0303
Root mean squared error           0.1519
Relative absolute error            26.4365 %
Root relative squared error       64.4699 %
Total Number of Instances         1509

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC     ROC Area   PRC Area   Class
0.987     0.247    0.985     0.987    0.986     0.751    0.940     0.993    negative
0.753     0.013    0.779     0.753    0.766     0.751    0.940     0.761    sick
Weighted Avg.   0.973    0.233    0.972     0.973    0.973     0.751    0.940     0.979

==== Confusion Matrix ====

      a      b  <-- classified as
1401   19 |   a = negative
   22   67 |   b = sick

```

## Percentage Split 66%

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B10-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  78
              age_1
              age_2
              age_3
              age_4
              age_5
              age_6
              age_7
              age_8
              age_9
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              TSH_2
              TSH_3
              TSH_4
              TSH_5
              TSH_6
              TSH_7
              TSH_8
              TSH_9
              T3_measured
              T3_1
              T3_2
              T3_3
              T3_4
              T3_5
              T3_6
              T3_7
              T3_8
              T3_9
              TT4_measured
              TT4_1
              TT4_2
              TT4_3
              TT4_4
              TT4_5
              TT4_6
              TT4_7
              TT4_8

```

```

TT4_9
T4U_measured
T4U_1
T4U_2
T4U_3
T4U_4
T4U_5
T4U_6
T4U_7
T4U_8
T4U_9
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
FTI_5
FTI_6
FTI_7
FTI_8
FTI_9
TBG_measured
TBG
referral_source
Class
Test mode: split 66.0% train, remainder test

==== Classifier model (full training set) ====

PART decision list
-----
T3_1 = '(1.15-inf)' AND
TT4_9 = '(-inf-149.5]' AND
TT4_1 = '(72.5-inf)': negative (2774.73)

T3_1 = '(1.15-inf)' AND
referral_source = other AND
FTI_9 = '(-inf-147.5)': negative (247.45/0.6)

T3_1 = '(1.15-inf)' AND
on_thyroxine = f AND
T3_9 = '(-inf-2.85]' AND
TSH_2 = '(0.285-inf)': negative (131.1/4.7)

referral_source = other AND
age_9 = '(-inf-75.5]' AND
T3_measured = f: negative (98.87/0.28)

TT4_9 = '(149.5-inf)' AND
T3_9 = '(2.85-inf)': negative (112.69)

referral_source = other AND
sick = f AND
age_9 = '(-inf-75.5]' AND
T4U_7 = '(1.055-inf)': negative (28.34)

T3_measured = t AND
T3_4 = '(1.85-inf)' AND
on_thyroxine = t: negative (21.66)

T3_measured = t AND
T3_9 = '(-inf-2.85]' AND
TT4_1 = '(72.5-inf)' AND
TT4_9 = '(-inf-149.5]' AND
TSH_measured = t AND
referral_source = SVI AND
FTI_9 = '(-inf-147.5)': sick (121.15/1.0)

T3_measured = f: negative (13.53/1.41)

T3_9 = '(-inf-2.85]' AND
sick = t AND
T3_1 = '(-inf-1.15)': sick (12.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
T4U_measured = t AND
T4U_8 = '(-inf-1.115]' AND
T4U_4 = '(0.935-inf)' AND
age_1 = '(25.5-inf)' AND
age_8 = '(-inf-70.5]' AND
age_6 = '(59.5-inf)': negative (12.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
T4U_measured = t AND

```

```

T4U_8 = '(-inf-1.115]' AND
psych = f AND
TSH_9 = '(6.15-inf)' AND
TT4_1 = '(-inf-72.5]' AND
on_thyroxine = f: negative (34.83/5.55)

T3_9 = '(-inf-2.85]' AND
sick = f AND
T4U_measured = t AND
T4U_8 = '(-inf-1.115]' AND
psych = f AND
TSH_measured = t AND
TSH_8 = '(3.05-inf)' AND
age_5 = '(54.5-inf)' AND
on_thyroxine = f: sick (12.0/2.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
T4U_measured = f: negative (10.83/0.85)

T3_9 = '(-inf-2.85]' AND
T4U_8 = '(1.115-inf)': negative (7.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
psych = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
on_thyroxine = t AND
age_5 = '(54.5-inf)' AND
TSH_3 = '(-inf-0.655)': negative (6.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
psych = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
T4U_4 = '(-inf-0.935]' AND
T3_6 = '(-inf-2.05]' AND
T3_4 = '(-inf-1.85]' AND
T3_1 = '(-inf-1.15]' AND
age_1 = '(25.5-inf)' AND
TSH_1 = '(0.085-inf)' AND
TT4_8 = '(-inf-130.5]' AND
age_7 = '(64.5-inf)' AND
TSH_3 = '(0.655-inf)' AND
FTI_1 = '(78.5-inf)': sick (20.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
psych = f AND
T4U_1 = '(0.785-inf)' AND
TSH_8 = '(-inf-3.05]' AND
query_hyperthyroid = f AND
T4U_3 = '(-inf-0.895]' AND
TT4_9 = '(-inf-149.5]' AND
FTI_1 = '(78.5-inf)': negative (10.0/3.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
psych = f AND
T4U_1 = '(0.785-inf)' AND
T4U_3 = '(-inf-0.895)': sick (20.72)

T3_7 = '(2.25-inf)': negative (6.39)

psych = f AND
sick = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
T4U_7 = '(-inf-1.055]' AND
T3_2 = '(-inf-1.47]' AND
TT4_7 = '(-inf-119.5]' AND
FTI_7 = '(120.5-inf)' AND
TSH_2 = '(-inf-0.285)': negative (4.0)

psych = f AND
sick = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
TSH_3 = '(0.655-inf)' AND
query_hypothyroid = f AND
on_thyroxine = f AND
TT4_4 = '(-inf-96.5]' AND
age_1 = '(25.5-inf)' AND
age_4 = '(-inf-47.5)': negative (5.0/1.0)

```

```

psych = f AND
sick = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
TSH_3 = '(0.655-inf)' AND
FTI_6 = '(-inf-113.5]' AND
query_hypothyroid = f AND
on_thyroxine = f AND
age_4 = '(47.5-inf)' AND
T4U_1 = '(-inf-0.785]' AND
FTI_1 = '(-inf-78.5]': negative (5.0/1.0)

sick = f AND
psych = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
TSH_1 = '(-inf-0.085]' AND
age_9 = '(-inf-75.5]' AND
referral_source = SVI AND
T4U_1 = '(0.785-inf)': negative (4.0)

sick = f AND
T3_5 = '(-inf-1.95]' AND
psych = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
age_6 = '(-inf-59.5]' AND
FTI_2 = '(-inf-89.5]': sick (7.0)

T3_5 = '(-inf-1.95]' AND
psych = f AND
T4U_6 = '(-inf-1.015]' AND
query_on_thyroxine = f AND
sick = f AND
T3_1 = '(1.15-inf)': sick (6.0/1.0)

T3_5 = '(-inf-1.95]' AND
psych = f AND
T3_1 = '(1.15-inf)': negative (4.26)

T3_1 = '(-inf-1.15]' AND
psych = f AND
age_9 = '(75.5-inf)' AND
T4U_4 = '(0.935-inf)': negative (3.0/1.0)

TT4_8 = '(-inf-130.5]' AND
age_9 = '(-inf-75.5]' AND
psych = f AND
referral_source = other: negative (9.72/1.0)

T4U_1 = '(-inf-0.785]' AND
TSH_2 = '(0.285-inf)' AND
FTI_3 = '(95.5-inf)': sick (7.5/3.0)

T4U_1 = '(0.785-inf)': sick (6.0)

TSH_2 = '(-inf-0.285]': sick (5.13/1.0)

: negative (4.1/1.1)

Number of Rules : 33

Time taken to build model: 0.21 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0.01 seconds

==== Summary ===

Correctly Classified Instances      1256          97.9719 %
Incorrectly Classified Instances    26           2.0281 %
Kappa statistic                   0.801
Mean absolute error               0.0243
Root mean squared error           0.137
Relative absolute error           21.255 %
Root relative squared error      58.3848 %
Total Number of Instances         1282

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC     ROC Area   PRC Area   Class
0.994      0.253      0.984      0.994      0.989      0.804      0.961      0.996      negative
0.747      0.006      0.889      0.747      0.812      0.804      0.961      0.791      sick
Weighted Avg.   0.980      0.239      0.979      0.980      0.979      0.804      0.961      0.984

```

```
==== Confusion Matrix ====
      a     b  <-- classified as
1200    7 |     a = negative
      19   56 |     b = sick
```

## Leave one out fold

```
==== Run information ====
Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B10-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  78
              age_1
              age_2
              age_3
              age_4
              age_5
              age_6
              age_7
              age_8
              age_9
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              t131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              TSH_2
              TSH_3
              TSH_4
              TSH_5
              TSH_6
              TSH_7
              TSH_8
              TSH_9
              T3_measured
              T3_1
              T3_2
              T3_3
              T3_4
              T3_5
              T3_6
              T3_7
              T3_8
              T3_9
              TT4_measured
              TT4_1
              TT4_2
              TT4_3
              TT4_4
              TT4_5
              TT4_6
              TT4_7
              TT4_8
              TT4_9
              T4U_measured
              T4U_1
              T4U_2
              T4U_3
              T4U_4
              T4U_5
              T4U_6
              T4U_7
              T4U_8
              T4U_9
              FTI_measured
              FTI_1
              FTI_2
              FTI_3
              FTI_4
              FTI_5
```

```

FTI_6
FTI_7
FTI_8
FTI_9
TBG_measured
TBG
referral_source
Class
Test mode: 3772-fold cross-validation

==== Classifier model (full training set) ===

PART decision list
-----
T3_1 = '(1.15-inf)' AND
TT4_9 = '(-inf-149.5)' AND
TT4_1 = '(72.5-inf)': negative (2774.73)

T3_1 = '(1.15-inf)' AND
referral_source = other AND
FTI_9 = '(-inf-147.5)': negative (247.45/0.6)

T3_1 = '(1.15-inf)' AND
on_thyroxine = f AND
T3_9 = '(-inf-2.85)' AND
TSH_2 = '(0.285-inf)': negative (131.1/4.7)

referral_source = other AND
age_9 = '(-inf-75.5)' AND
T3_measured = f: negative (98.87/0.28)

TT4_9 = '(149.5-inf)' AND
T3_9 = '(2.85-inf)': negative (112.69)

referral_source = other AND
sick = f AND
age_9 = '(-inf-75.5)' AND
T4U_7 = '(1.055-inf)': negative (28.34)

T3_measured = t AND
T3_4 = '(1.85-inf)' AND
on_thyroxine = t: negative (21.66)

T3_measured = t AND
T3_9 = '(-inf-2.85)' AND
TT4_1 = '(72.5-inf)' AND
TT4_9 = '(-inf-149.5)' AND
TSH_measured = t AND
referral_source = SVI AND
FTI_9 = '(-inf-147.5)': sick (121.15/1.0)

T3_measured = f: negative (13.53/1.41)

T3_9 = '(-inf-2.85)' AND
sick = t AND
T3_1 = '(-inf-1.15)': sick (12.0)

T3_9 = '(-inf-2.85)' AND
sick = f AND
T4U_measured = t AND
T4U_8 = '(-inf-1.115)' AND
T4U_4 = '(0.935-inf)' AND
age_1 = '(25.5-inf)' AND
age_8 = '(-inf-70.5)' AND
age_6 = '(59.5-inf)': negative (12.0)

T3_9 = '(-inf-2.85)' AND
sick = f AND
T4U_measured = t AND
T4U_8 = '(-inf-1.115)' AND
psych = f AND
TSH_9 = '(6.15-inf)' AND
TT4_1 = '(-inf-72.5)' AND
on_thyroxine = f: negative (34.83/5.55)

T3_9 = '(-inf-2.85)' AND
sick = f AND
T4U_measured = t AND
T4U_8 = '(-inf-1.115)' AND
psych = f AND
TSH_measured = t AND
TSH_8 = '(3.05-inf)' AND
age_5 = '(54.5-inf)' AND
on_thyroxine = f: sick (12.0/2.0)

T3_9 = '(-inf-2.85)' AND

```

```

sick = f AND
T4U_measured = f: negative (10.83/0.85)

T3_9 = '(-inf-2.85]' AND
T4U_8 = '(1.115-inf)': negative (7.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
psych = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
on_thyroxine = t AND
age_5 = '(54.5-inf)' AND
TSH_3 = '(-inf-0.655)': negative (6.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
psych = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
T4U_4 = '(-inf-0.935]' AND
T3_6 = '(-inf-2.05]' AND
T3_4 = '(-inf-1.85]' AND
T3_1 = '(-inf-1.15]' AND
age_1 = '(25.5-inf)' AND
TSH_1 = '(0.085-inf)' AND
TT4_8 = '(-inf-130.5]' AND
age_7 = '(64.5-inf)' AND
TSH_3 = '(0.655-inf)' AND
FTI_1 = '(78.5-inf)': sick (20.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
psych = f AND
T4U_1 = '(0.785-inf)' AND
TSH_8 = '(-inf-3.05]' AND
query_hyperthyroid = f AND
T4U_3 = '(-inf-0.895]' AND
T4U_9 = '(-inf-149.5]' AND
FTI_1 = '(78.5-inf)': negative (10.0/3.0)

T3_9 = '(-inf-2.85]' AND
sick = f AND
psych = f AND
T4U_1 = '(0.785-inf)' AND
T4U_3 = '(-inf-0.895)': sick (20.72)

T3_7 = '(2.25-inf)': negative (6.39)

psych = f AND
sick = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
T4U_7 = '(-inf-1.055]' AND
T3_2 = '(-inf-1.47]' AND
TT4_7 = '(-inf-119.5]' AND
FTI_7 = '(120.5-inf)' AND
TSH_2 = '(-inf-0.285)': negative (4.0)

psych = f AND
sick = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
TSH_3 = '(0.655-inf)' AND
query_hypothyroid = f AND
on_thyroxine = f AND
TT4_4 = '(-inf-96.5]' AND
age_1 = '(25.5-inf)' AND
age_4 = '(-inf-47.5)': negative (5.0/1.0)

psych = f AND
sick = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
TSH_3 = '(0.655-inf)' AND
FTI_6 = '(-inf-113.5]' AND
query_hypothyroid = f AND
on_thyroxine = f AND
age_4 = '(47.5-inf)' AND
T4U_1 = '(-inf-0.785]' AND
FTI_1 = '(-inf-78.5)': negative (5.0/1.0)

sick = f AND
psych = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
TSH_1 = '(-inf-0.085]' AND

```

```

age_9 = '(-inf-75.5]' AND
referral_source = SVI AND
T4U_1 = '(0.785-inf)': negative (4.0)

sick = f AND
T3_5 = '(-inf-1.95]' AND
psych = f AND
TSH_measured = t AND
query_hyperthyroid = f AND
age_6 = '(-inf-59.5]' AND
FTI_2 = '(-inf-89.5]': sick (7.0)

T3_5 = '(-inf-1.95]' AND
psych = f AND
T4U_6 = '(-inf-1.015]' AND
query_on_thyroxine = f AND
sick = f AND
T3_1 = '(1.15-inf)': sick (6.0/1.0)

T3_5 = '(-inf-1.95]' AND
psych = f AND
T3_1 = '(1.15-inf)': negative (4.26)

T3_1 = '(-inf-1.15]' AND
psych = f AND
age_9 = '(75.5-inf)' AND
T4U_4 = '(0.935-inf)': negative (3.0/1.0)

TT4_8 = '(-inf-130.5]' AND
age_9 = '(-inf-75.5]' AND
psych = f AND
referral_source = other: negative (9.72/1.0)

T4U_1 = '(-inf-0.785]' AND
TSH_2 = '(0.285-inf)' AND
FTI_3 = '(95.5-inf)': sick (7.5/3.0)

T4U_1 = '(0.785-inf)': sick (6.0)

TSH_2 = '(-inf-0.285)': sick (5.13/1.0)

: negative (4.1/1.1)

Number of Rules : 33

Time taken to build model: 0.2 seconds

==== Stratified cross-validation ====
==== Summary ====

    Correctly Classified Instances      3680          97.561 %
    Incorrectly Classified Instances     92           2.439 %
    Kappa statistic                      0.7879
    Mean absolute error                  0.0264
    Root mean squared error              0.1414
    Relative absolute error              22.9381 %
    Root relative squared error         58.9765 %
    Total Number of Instances            3772

==== Detailed Accuracy By Class ====

    TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
    0.987     0.199     0.987     0.987     0.987     0.788     0.963     0.996     negative
    0.801     0.013     0.801     0.801     0.801     0.788     0.963     0.817     sick
    Weighted Avg.   0.976     0.188     0.976     0.976     0.976     0.788     0.963     0.985

==== Confusion Matrix ====

    a      b  <-- classified as
3495   46 |     a = negative
    46  185 |     b = sick

```

### 3.2.4 40 bins

#### Cross Validation 5 folds

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B40-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  258

```

```

[list of attributes omitted]
Test mode: 5-fold cross-validation

==== Classifier model (full training set) ====

PART decision list
-----

T3_4 = '(1.15-inf)' AND
FTI_38 = '(-inf-165.5)' AND
TT4_2 = '(62.5-inf)': negative (3118.35/1.0)

referral_source = other AND
age_35 = '(-inf-72.5)' AND
sick = f AND
TSH_measured = t AND
T4U_25 = '(1.025-inf)': negative (78.35)

T3_measured = f: negative (115.58/3.0)

T3_4 = '(1.15-inf)' AND
T3_19 = '(2.65-inf)': negative (58.57)

T3_4 = '(1.15-inf)' AND
TT4_36 = '(-inf-150.5)' AND
T4U_12 = '(-inf-0.895)': negative (44.38)

TT4_1 = '(-inf-46.5)' AND
age_38 = '(-inf-78.5)' AND
TSH_38 = '(12.55-inf)': negative (38.46/1.0)

TT4_39 = '(-inf-192.5)' AND
referral_source = SVI AND
TSH_39 = '(-inf-29.5)' AND
TSH_38 = '(-inf-12.55)' AND
FTI_37 = '(-inf-153.5)' AND
TT4_1 = '(46.5-inf)' AND
T4U_33 = '(-inf-1.135)' AND
TT4_2 = '(62.5-inf)': sick (141.18/0.25)

sick = f AND
TT4_39 = '(192.5-inf)': negative (8.05)

sick = f AND
age_1 = '(16.5-inf)' AND
T3_18 = '(-inf-2.55)' AND
T4U_measured = f: negative (7.72/0.9)

sick = t: sick (7.13/0.01)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
T3_18 = '(-inf-2.55)' AND
age_39 = '(82.5-inf)' AND
TT4_32 = '(-inf-132.5)': sick (6.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
T3_18 = '(-inf-2.55)' AND
referral_source = other AND
FTI_2 = '(68.5-inf)' AND
age_37 = '(-inf-76.5)' AND
TSH_34 = '(-inf-4.05)' AND
T3_1 = '(0.55-inf)' AND
T3_9 = '(-inf-1.65)' AND
age_33 = '(-inf-70.5)': negative (21.67/1.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
psych = f AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
query_hyperthyroid = f AND
I131_treatment = f AND
T3_18 = '(-inf-2.55)' AND
FTI_35 = '(-inf-141.5)' AND
T3_10 = '(-inf-1.75)' AND
TT4_5 = '(75.5-inf)': sick (12.0)

age_1 = '(16.5-inf)' AND

```

```

T4U_34 = '(-inf-1.155)' AND
psych = f AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
query_hyperthyroid = f AND
I131_treatment = f AND
T3_18 = '(-inf-2.55]' AND
T4U_13 = '(0.905-inf)' AND
T4U_22 = '(-inf-0.995)' AND
T3_5 = '(1.25-inf)': negative (15.0/1.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
psych = f AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
query_hyperthyroid = f AND
I131_treatment = f AND
T3_18 = '(-inf-2.55]' AND
query_on_thyroxine = f AND
age_26 = '(-inf-60.5]' AND
age_8 = '(32.5-inf)' AND
T3_9 = '(-inf-1.65)' AND
TT4_10 = '(-inf-87.5)': sick (18.0/3.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
psych = f AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_27 = '(-inf-1.045)' AND
query_hyperthyroid = f AND
I131_treatment = f AND
T3_18 = '(-inf-2.55]' AND
query_on_thyroxine = f AND
T3_4 = '(1.15-inf)' AND
on_thyroxine = f AND
TSH_7 = '(-inf-0.205)': sick (11.0)

psych = f AND
query_on_thyroxine = f AND
T3_18 = '(-inf-2.55]' AND
TSH_4 = '(0.085-inf)' AND
query_hyperthyroid = f AND
TT4_21 = '(106.5-inf)': negative (13.02)

T4U_34 = '(-inf-1.155)' AND
TT4_37 = '(-inf-159.5]' AND
TSH_4 = '(0.085-inf)' AND
TT4_2 = '(62.5-inf)': sick (14.32/1.4)

TSH_37 = '(-inf-8.25]' AND
T4U_1 = '(0.685-inf)' AND
on_thyroxine = f: sick (17.14/2.05)

sex = F: negative (17.0/0.0)

TT4_3 = '(68.5-inf)': negative (4.72)
: sick (4.38/1.35)

Number of Rules : 22

Time taken to build model: 0.59 seconds

==== Stratified cross-validation ====
==== Summary ====

  Correctly Classified Instances      3717          98.5419 %
  Incorrectly Classified Instances     55           1.4581 %
  Kappa statistic                   0.8724
  Mean absolute error                0.0176
  Root mean squared error            0.1169
  Relative absolute error             15.231 %
  Root relative squared error        48.7703 %
  Total Number of Instances         3772

==== Detailed Accuracy By Class ====

          TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
          0.993     0.126     0.992     0.993     0.992     0.872     0.954     0.993     negative
          0.874     0.007     0.886     0.874     0.880     0.872     0.954     0.866     sick
  Weighted Avg.       0.985     0.118     0.985     0.985     0.985     0.872     0.954     0.985

```

```
==== Confusion Matrix ====
      a     b  <-- classified as
3515   26 |     a = negative
      29  202 |     b = sick
```

## Cross Validation 10 folds

```
==== Run information ====
Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B40-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  258
              [list of attributes omitted]
Test mode:   10-fold cross-validation

==== Classifier model (full training set) ====
PART decision list
-----
T3_4 = '(1.15-inf)' AND
FTI_38 = '(-inf-165.5)' AND
TT4_2 = '(62.5-inf)': negative (3118.35/1.0)

referral_source = other AND
age_35 = '(-inf-72.5)' AND
sick = f AND
TSH_measured = t AND
T4U_25 = '(1.025-inf)': negative (78.35)

T3_measured = f: negative (115.58/3.0)

T3_4 = '(1.15-inf)' AND
T3_19 = '(2.65-inf)': negative (58.57)

T3_4 = '(1.15-inf)' AND
TT4_36 = '(-inf-150.5)' AND
T4U_12 = '(-inf-0.895)': negative (44.38)

TT4_1 = '(-inf-46.5)' AND
age_38 = '(-inf-78.5)' AND
TSH_38 = '(12.55-inf)': negative (38.46/1.0)

TT4_39 = '(-inf-192.5)' AND
referral_source = SVI AND
TSH_39 = '(-inf-29.5)' AND
TSH_38 = '(-inf-12.55)' AND
FTI_37 = '(-inf-153.5)' AND
TT4_1 = '(46.5-inf)' AND
T4U_33 = '(-inf-1.135)' AND
TT4_2 = '(62.5-inf)': sick (141.18/0.25)

sick = f AND
TT4_39 = '(192.5-inf)': negative (8.05)

sick = f AND
age_1 = '(16.5-inf)' AND
T3_18 = '(-inf-2.55)' AND
T4U_measured = f: negative (7.72/0.9)

sick = t: sick (7.13/0.01)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
T3_18 = '(-inf-2.55)' AND
age_39 = '(82.5-inf)' AND
TT4_32 = '(-inf-132.5)': sick (6.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
T3_18 = '(-inf-2.55)' AND
referral_source = other AND
FTI_2 = '(68.5-inf)' AND
age_37 = '(-inf-76.5)' AND
TSH_34 = '(-inf-4.05)' AND
T3_1 = '(0.55-inf)' AND
```

```

T3_9 = '(-inf-1.65)' AND
age_33 = '(-inf-70.5)': negative (21.67/1.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
psych = f AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
query_hyperthyroid = f AND
I131_treatment = f AND
T3_18 = '(-inf-2.55)' AND
FTI_35 = '(-inf-141.5)' AND
T3_10 = '(-inf-1.75)' AND
TT4_5 = '(75.5-inf)': sick (12.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
psych = f AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
query_hyperthyroid = f AND
I131_treatment = f AND
T3_18 = '(-inf-2.55)' AND
T4U_13 = '(0.905-inf)' AND
T4U_22 = '(-inf-0.995)' AND
T3_5 = '(1.25-inf)': negative (15.0/1.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
psych = f AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
query_hyperthyroid = f AND
I131_treatment = f AND
T3_18 = '(-inf-2.55)' AND
query_on_thyroxine = f AND
age_26 = '(-inf-60.5)' AND
age_8 = '(32.5-inf)' AND
T3_9 = '(-inf-1.65)' AND
TT4_10 = '(-inf-87.5)': sick (18.0/3.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
psych = f AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_27 = '(-inf-1.045)' AND
query_hyperthyroid = f AND
I131_treatment = f AND
T3_18 = '(-inf-2.55)' AND
query_on_thyroxine = f AND
T3_4 = '(1.15-inf)' AND
on_thyroxine = f AND
TSH_7 = '(-inf-0.205)': sick (11.0)

psych = f AND
query_on_thyroxine = f AND
T3_18 = '(-inf-2.55)' AND
TSH_4 = '(0.085-inf)' AND
query_hyperthyroid = f AND
TT4_21 = '(106.5-inf)': negative (13.02)

T4U_34 = '(-inf-1.155)' AND
TT4_37 = '(-inf-159.5)' AND
TSH_4 = '(0.085-inf)' AND
TT4_2 = '(62.5-inf)': sick (14.32/1.4)

TSH_37 = '(-inf-8.25)' AND
T4U_1 = '(0.685-inf)' AND
on_thyroxine = f: sick (17.14/2.05)

sex = F: negative (17.0/0.0)

TT4_3 = '(68.5-inf)': negative (4.72)

: sick (4.38/1.35)

Number of Rules : 22

```

```

Time taken to build model: 0.56 seconds

== Stratified cross-validation ==
== Summary ==

```

```

Correctly Classified Instances      3707          98.2768 %
Incorrectly Classified Instances   65           1.7232 %
Kappa statistic                   0.8454
Mean absolute error               0.0194
Root mean squared error           0.1256
Relative absolute error           16.8028 %
Root relative squared error      52.3969 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC     ROC Area  PRC Area  Class
          0.993    0.173    0.989    0.993    0.991    0.846    0.933    0.988  negative
          0.827    0.007    0.884    0.827    0.855    0.846    0.933    0.852   sick
Weighted Avg.      0.983    0.163    0.982    0.983    0.982    0.846    0.933    0.979

==== Confusion Matrix ====

      a     b  <-- classified as
3516  25 |   a = negative
  40  191 |   b = sick

```

## Cross Validation 15 folds

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B40-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  258
              [list of attributes omitted]
Test mode:   15-fold cross-validation

==== Classifier model (full training set) ===

PART decision list
-----
T3_4 = '(1.15-inf)' AND
FTI_38 = '(-inf-165.5)' AND
TT4_2 = '(62.5-inf)': negative (3118.35/1.0)

referral_source = other AND
age_35 = '(-inf-72.5)' AND
sick = f AND
TSH_measured = t AND
T4U_25 = '(1.025-inf)': negative (78.35)

T3_measured = f: negative (115.58/3.0)

T3_4 = '(1.15-inf)' AND
T3_19 = '(2.65-inf)': negative (58.57)

T3_4 = '(1.15-inf)' AND
TT4_36 = '(-inf-150.5)' AND
T4U_12 = '(-inf-0.895)': negative (44.38)

TT4_1 = '(-inf-46.5)' AND
age_38 = '(-inf-78.5)' AND
TSH_38 = '(12.55-inf)': negative (38.46/1.0)

TT4_39 = '(-inf-192.5)' AND
referral_source = SVI AND
TSH_39 = '(-inf-29.5)' AND
TSH_38 = '(-inf-12.55)' AND
FTI_37 = '(-inf-153.5)' AND
TT4_1 = '(46.5-inf)' AND
T4U_33 = '(-inf-1.135)' AND
TT4_2 = '(62.5-inf)': sick (141.18/0.25)

sick = f AND
TT4_39 = '(192.5-inf)': negative (8.05)

sick = f AND
age_1 = '(16.5-inf)' AND
T3_18 = '(-inf-2.55)' AND
T4U_measured = f: negative (7.72/0.9)

sick = t: sick (7.13/0.01)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
T4U_32 = '(-inf-1.115)' AND

```

```

T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
T3_18 = '(-inf-2.55)' AND
age_39 = '(82.5-inf)' AND
TT4_32 = '(-inf-132.5)': sick (6.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
T3_18 = '(-inf-2.55]' AND
referral_source = other AND
FTI_2 = '(68.5-inf)' AND
age_37 = '(-inf-76.5)' AND
TSH_34 = '(-inf-4.05)' AND
T3_1 = '(0.55-inf)' AND
T3_9 = '(-inf-1.65)' AND
age_33 = '(-inf-70.5)': negative (21.67/1.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
psych = f AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
query_hyperthyroid = f AND
I131_treatment = f AND
T3_18 = '(-inf-2.55]' AND
FTI_35 = '(-inf-141.5)' AND
T3_10 = '(-inf-1.75)' AND
TT4_5 = '(75.5-inf)': sick (12.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
psych = f AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
query_hyperthyroid = f AND
I131_treatment = f AND
T3_18 = '(-inf-2.55)' AND
T4U_13 = '(0.905-inf)' AND
T4U_22 = '(-inf-0.995)' AND
T3_5 = '(1.25-inf)': negative (15.0/1.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
psych = f AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
query_hyperthyroid = f AND
I131_treatment = f AND
T3_18 = '(-inf-2.55)' AND
query_on_thyroxine = f AND
age_26 = '(-inf-60.5)' AND
age_8 = '(32.5-inf)' AND
T3_9 = '(-inf-1.65)' AND
TT4_10 = '(-inf-87.5)': sick (18.0/3.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
psych = f AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_27 = '(-inf-1.045)' AND
query_hyperthyroid = f AND
I131_treatment = f AND
T3_18 = '(-inf-2.55)' AND
query_on_thyroxine = f AND
T3_4 = '(1.15-inf)' AND
on_thyroxine = f AND
TSH_7 = '(-inf-0.205)': sick (11.0)

psych = f AND
query_on_thyroxine = f AND
T3_18 = '(-inf-2.55)' AND
TSH_4 = '(0.085-inf)' AND
query_hyperthyroid = f AND
TT4_21 = '(106.5-inf)': negative (13.02)

T4U_34 = '(-inf-1.155)' AND
TT4_37 = '(-inf-159.5)' AND
TSH_4 = '(0.085-inf)' AND
TT4_2 = '(62.5-inf)': sick (14.32/1.4)

```

```

TSH_37 = '(-inf-8.25]' AND
T4U_1 = '(0.685-inf)' AND
on_thyroxine = f: sick (17.14/2.05)

sex = F: negative (17.0/0.0)

TT4_3 = '(68.5-inf)': negative (4.72)

: sick (4.38/1.35)

Number of Rules : 22

Time taken to build model: 0.58 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3695          97.9586 %
Incorrectly Classified Instances   77           2.0414 %
Kappa statistic                   0.8242
Mean absolute error               0.0221
Root mean squared error          0.1337
Relative absolute error          19.1775 %
Root relative squared error     55.7507 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC    ROC Area  PRC Area  Class
0.988     0.156     0.990     0.988     0.989     0.824     0.953     0.991  negative
0.844     0.012     0.826     0.844     0.835     0.824     0.953     0.816   sick
Weighted Avg.  0.980     0.147     0.980     0.980     0.980     0.824     0.953     0.980

==== Confusion Matrix ====

      a      b  <-- classified as
3500   41 |   a = negative
      36  195 |   b = sick

```

## Percentage Split 50%

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B40-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  258
              [list of attributes omitted]
Test mode:   split 50.0% train, remainder test

==== Classifier model (full training set) ====

PART decision list
-----
T3_4 = '(1.15-inf)' AND
FTI_38 = '(-inf-165.5]' AND
TT4_2 = '(62.5-inf)': negative (3118.35/1.0)

referral_source = other AND
age_35 = '(-inf-72.5]' AND
sick = f AND
TSH_measured = t AND
T4U_25 = '(1.025-inf)': negative (78.35)

T3_measured = f: negative (115.58/3.0)

T3_4 = '(1.15-inf)' AND
T3_19 = '(2.65-inf)': negative (58.57)

T3_4 = '(1.15-inf)' AND
TT4_36 = '(-inf-150.5]' AND
T4U_12 = '(-inf-0.895)': negative (44.38)

TT4_1 = '(-inf-46.5]' AND
age_38 = '(-inf-78.5]' AND
TSH_38 = '(12.55-inf)': negative (38.46/1.0)

TT4_39 = '(-inf-192.5]' AND
referral_source = SVI AND
TSH_39 = '(-inf-29.5]' AND
TSH_38 = '(-inf-12.55]' AND
FTI_37 = '(-inf-153.5]' AND

```

```

TT4_1 = '(46.5-inf)' AND
T4U_33 = '(-inf-1.135)' AND
TT4_2 = '(62.5-inf)': sick (141.18/0.25)

sick = f AND
TT4_39 = '(192.5-inf)': negative (8.05)

sick = f AND
age_1 = '(16.5-inf)' AND
T3_18 = '(-inf-2.55]' AND
T4U_measured = f: negative (7.72/0.9)

sick = t: sick (7.13/0.01)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
T3_18 = '(-inf-2.55]' AND
age_39 = '(82.5-inf)' AND
TT4_32 = '(-inf-132.5)': sick (6.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
T3_18 = '(-inf-2.55]' AND
referral_source = other AND
FTI_2 = '(68.5-inf)' AND
age_37 = '(-inf-76.5)' AND
TSH_34 = '(-inf-4.05]' AND
T3_1 = '(0.55-inf)' AND
T3_9 = '(-inf-1.65)' AND
age_33 = '(-inf-70.5)': negative (21.67/1.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
psych = f AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
query_hyperthyroid = f AND
I131_treatment = f AND
T3_18 = '(-inf-2.55]' AND
FTI_35 = '(-inf-141.5)' AND
T3_10 = '(-inf-1.75]' AND
TT4_5 = '(75.5-inf)': sick (12.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
psych = f AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
query_hyperthyroid = f AND
I131_treatment = f AND
T3_18 = '(-inf-2.55]' AND
T4U_13 = '(0.905-inf)' AND
T4U_22 = '(-inf-0.995)' AND
T3_5 = '(1.25-inf)': negative (15.0/1.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
psych = f AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
query_hyperthyroid = f AND
I131_treatment = f AND
T3_18 = '(-inf-2.55]' AND
query_on_thyroxine = f AND
age_26 = '(-inf-60.5)' AND
age_8 = '(32.5-inf)' AND
T3_9 = '(-inf-1.65)' AND
TT4_10 = '(-inf-87.5)': sick (18.0/3.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
psych = f AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_27 = '(-inf-1.045)' AND
query_hyperthyroid = f AND
I131_treatment = f AND
T3_18 = '(-inf-2.55]' AND

```

```

query_on_thyroxine = f AND
T3_4 = '(1.15-inf)' AND
on_thyroxine = f AND
TSH_7 = '(-inf-0.205)': sick (11.0)

psych = f AND
query_on_thyroxine = f AND
T3_18 = '(-inf-2.55]' AND
TSH_4 = '(0.085-inf)' AND
query_hypothyroid = f AND
TT4_21 = '(106.5-inf)': negative (13.02)

T4U_34 = '(-inf-1.155]' AND
TT4_37 = '(-inf-159.5]' AND
TSH_4 = '(0.085-inf)' AND
TT4_2 = '(62.5-inf)': sick (14.32/1.4)

TSH_37 = '(-inf-8.25]' AND
T4U_1 = '(0.685-inf)' AND
on_thyroxine = f: sick (17.14/2.05)

sex = F: negative (17.0/0.0)

TT4_3 = '(68.5-inf)': negative (4.72)

: sick (4.38/1.35)

Number of Rules : 22

Time taken to build model: 0.64 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0.01 seconds

==== Summary ===

Correctly Classified Instances      1860           98.6214 %
Incorrectly Classified Instances   26            1.3786 %
Kappa statistic                   0.8745
Mean absolute error               0.0179
Root mean squared error          0.1114
Relative absolute error          15.4886 %
Root relative squared error     46.1723 %
Total Number of Instances        1886

==== Detailed Accuracy By Class ===

          TP Rate  FP Rate  Precision  Recall  F-Measure  MCC    ROC Area  PRC Area  Class
          0.997    0.171    0.989    0.997    0.993    0.877    0.963    0.994    negative
          0.829    0.003    0.942    0.829    0.882    0.877    0.963    0.912    sick
Weighted Avg.    0.986    0.161    0.986    0.986    0.986    0.877    0.963    0.989

==== Confusion Matrix ===

      a      b  <-- classified as
1763    6 |      a = negative
      20    97 |      b = sick

```

## Percentage Split 60%

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B40-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  258
              [list of attributes omitted]
Test mode:   split 60.0% train, remainder test

==== Classifier model (full training set) ===

PART decision list
-----
T3_4 = '(1.15-inf)' AND
FTI_38 = '(-inf-165.5]' AND
TT4_2 = '(62.5-inf)': negative (3118.35/1.0)

referral_source = other AND
age_35 = '(-inf-72.5]' AND
sick = f AND
TSH_measured = t AND

```

```

T4U_25 = '(1.025-inf)': negative (78.35)
T3_measured = f: negative (115.58/3.0)

T3_4 = '(1.15-inf)' AND
T3_19 = '(2.65-inf)': negative (58.57)

T3_4 = '(1.15-inf)' AND
TT4_36 = '(-inf-150.5)' AND
T4U_12 = '(-inf-0.895)': negative (44.38)

TT4_1 = '(-inf-46.5]' AND
age_38 = '(-inf-78.5]' AND
TSH_38 = '(12.55-inf)': negative (38.46/1.0)

TT4_39 = '(-inf-192.5]' AND
referral_source = SVI AND
TSH_39 = '(-inf-29.5]' AND
TSH_38 = '(-inf-12.55]' AND
FTI_37 = '(-inf-153.5)' AND
TT4_1 = '(46.5-inf)' AND
T4U_33 = '(-inf-1.135)' AND
TT4_2 = '(62.5-inf)': sick (141.18/0.25)

sick = f AND
TT4_39 = '(192.5-inf)': negative (8.05)

sick = f AND
age_1 = '(16.5-inf)' AND
T3_18 = '(-inf-2.55]' AND
T4U_measured = f: negative (7.72/0.9)

sick = t: sick (7.13/0.01)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
T3_18 = '(-inf-2.55]' AND
age_39 = '(82.5-inf)' AND
TT4_32 = '(-inf-132.5)': sick (6.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
T3_18 = '(-inf-2.55]' AND
referral_source = other AND
FTI_2 = '(68.5-inf)' AND
age_37 = '(-inf-76.5)' AND
TSH_34 = '(-inf-4.05)' AND
T3_1 = '(0.55-inf)' AND
T3_9 = '(-inf-1.65)' AND
age_33 = '(-inf-70.5)': negative (21.67/1.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
psych = f AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
query_hyperthyroid = f AND
I131_treatment = f AND
T3_18 = '(-inf-2.55)' AND
FTI_35 = '(-inf-141.5)' AND
T3_10 = '(-inf-1.75)' AND
TT4_5 = '(75.5-inf)': sick (12.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
psych = f AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
query_hyperthyroid = f AND
I131_treatment = f AND
T3_18 = '(-inf-2.55)' AND
T4U_13 = '(0.905-inf)' AND
T4U_22 = '(-inf-0.995)' AND
T3_5 = '(1.25-inf)': negative (15.0/1.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
psych = f AND
T4U_32 = '(-inf-1.115)' AND

```

```

T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
query_hyperthyroid = f AND
I131_treatment = f AND
T3_18 = '(-inf-2.55]' AND
query_on_thyroxine = f AND
age_26 = '(-inf-60.5]' AND
age_8 = '(32.5-inf)' AND
T3_9 = '(-inf-1.65)' AND
TT4_10 = '(-inf-87.5)': sick (18.0/3.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
psych = f AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_27 = '(-inf-1.045)' AND
query_hyperthyroid = f AND
I131_treatment = f AND
T3_18 = '(-inf-2.55]' AND
query_on_thyroxine = f AND
T3_4 = '(1.15-inf)' AND
on_thyroxine = f AND
TSH_7 = '(-inf-0.205)': sick (11.0)

psych = f AND
query_on_thyroxine = f AND
T3_18 = '(-inf-2.55]' AND
TSH_4 = '(0.085-inf)' AND
query_hyperthyroid = f AND
TT4_21 = '(106.5-inf)': negative (13.02)

T4U_34 = '(-inf-1.155)' AND
TT4_37 = '(-inf-159.5)' AND
TSH_4 = '(0.085-inf)' AND
TT4_2 = '(62.5-inf)': sick (14.32/1.4)

TSH_37 = '(-inf-8.25)' AND
T4U_1 = '(0.685-inf)' AND
on_thyroxine = f: sick (17.14/2.05)

sex = F: negative (17.0/0.0)

TT4_3 = '(68.5-inf)': negative (4.72)

: sick (4.38/1.35)

Number of Rules : 22

Time taken to build model: 0.63 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0.07 seconds

==== Summary ===

    Correctly Classified Instances      1480          98.0782 %
    Incorrectly Classified Instances   29           1.9218 %
    Kappa statistic                   0.8222
    Mean absolute error              0.0223
    Root mean squared error          0.1363
    Relative absolute error          19.4713 %
    Root relative squared error     57.8675 %
    Total Number of Instances        1509

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.992     0.191     0.988     0.992     0.990     0.823     0.935     0.989     negative
      0.809     0.008     0.857     0.809     0.832     0.823     0.935     0.791     sick
Weighted Avg.      0.981     0.180     0.980     0.981     0.981     0.823     0.935     0.977

==== Confusion Matrix ===

      a      b  <-- classified as
1408    12 |      a = negative
      17    72 |      b = sick

```

## Percentage Split 66%

```
==== Run information ===
```

```

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B40-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  258
              [list of attributes omitted]
Test mode:   split 66.0% train, remainder test

==== Classifier model (full training set) ====

PART decision list
-----

T3_4 = '(1.15-inf)' AND
FTI_38 = '(-inf-165.5)' AND
TT4_2 = '(62.5-inf)': negative (3118.35/1.0)

referral_source = other AND
age_35 = '(-inf-72.5)' AND
sick = f AND
TSH_measured = t AND
T4U_25 = '(1.025-inf)': negative (78.35)

T3_measured = f: negative (115.58/3.0)

T3_4 = '(1.15-inf)' AND
T3_19 = '(2.65-inf)': negative (58.57)

T3_4 = '(1.15-inf)' AND
TT4_36 = '(-inf-150.5)' AND
T4U_12 = '(-inf-0.895)': negative (44.38)

TT4_1 = '(-inf-46.5)' AND
age_38 = '(-inf-78.5)' AND
TSH_38 = '(12.55-inf)': negative (38.46/1.0)

TT4_39 = '(-inf-192.5)' AND
referral_source = SVI AND
TSH_39 = '(-inf-29.5)' AND
TSH_38 = '(-inf-12.55)' AND
FTI_37 = '(-inf-153.5)' AND
TT4_1 = '(46.5-inf)' AND
T4U_33 = '(-inf-1.135)' AND
TT4_2 = '(62.5-inf)': sick (141.18/0.25)

sick = f AND
TT4_39 = '(192.5-inf)': negative (8.05)

sick = f AND
age_1 = '(16.5-inf)' AND
T3_18 = '(-inf-2.55)' AND
T4U_measured = f: negative (7.72/0.9)

sick = t: sick (7.13/0.01)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
T3_18 = '(-inf-2.55)' AND
age_39 = '(82.5-inf)' AND
TT4_32 = '(-inf-132.5)': sick (6.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
T3_18 = '(-inf-2.55)' AND
referral_source = other AND
FTI_2 = '(68.5-inf)' AND
age_37 = '(-inf-76.5)' AND
TSH_34 = '(-inf-4.05)' AND
T3_1 = '(0.55-inf)' AND
T3_9 = '(-inf-1.65)' AND
age_33 = '(-inf-70.5)': negative (21.67/1.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
psych = f AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
query_hyperthyroid = f AND
I131_treatment = f AND
T3_18 = '(-inf-2.55)' AND
FTI_35 = '(-inf-141.5)' AND

```

```

T3_10 = '(-inf-1.75]' AND
TT4_5 = '(75.5-inf)': sick (12.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
psych = f AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
query_hyperthyroid = f AND
I131_treatment = f AND
T3_18 = '(-inf-2.55]' AND
T4U_13 = '(0.905-inf)' AND
T4U_22 = '(-inf-0.995)' AND
T3_5 = '(1.25-inf)': negative (15.0/1.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
psych = f AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
query_hyperthyroid = f AND
I131_treatment = f AND
T3_18 = '(-inf-2.55]' AND
query_on_thyroxine = f AND
age_26 = '(-inf-60.5]' AND
age_8 = '(32.5-inf)' AND
T3_9 = '(-inf-1.65)' AND
TT4_10 = '(-inf-87.5)': sick (18.0/3.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
psych = f AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_27 = '(-inf-1.045)' AND
query_hyperthyroid = f AND
I131_treatment = f AND
T3_18 = '(-inf-2.55]' AND
query_on_thyroxine = f AND
T3_4 = '(1.15-inf)' AND
on_thyroxine = f AND
TSH_7 = '(-inf-0.205)': sick (11.0)

psych = f AND
query_on_thyroxine = f AND
T3_18 = '(-inf-2.55]' AND
TSH_4 = '(0.085-inf)' AND
query_hyperthyroid = f AND
TT4_21 = '(106.5-inf)': negative (13.02)

T4U_34 = '(-inf-1.155)' AND
TT4_37 = '(-inf-159.5)' AND
TSH_4 = '(0.085-inf)' AND
TT4_2 = '(62.5-inf)': sick (14.32/1.4)

TSH_37 = '(-inf-8.25]' AND
T4U_1 = '(0.685-inf)' AND
on_thyroxine = f: sick (17.14/2.05)

sex = F: negative (17.0/0.0)

TT4_3 = '(68.5-inf)': negative (4.72)

: sick (4.38/1.35)

Number of Rules : 22

Time taken to build model: 0.64 seconds
== Evaluation on test split ==
Time taken to test model on test split: 0.02 seconds
== Summary ==
Correctly Classified Instances      1260          98.2839 %
Incorrectly Classified Instances    22           1.7161 %
Kappa statistic                      0.8293
Mean absolute error                  0.0207
Root mean squared error              0.1216
Relative absolute error              18.1788 %
Root relative squared error         51.8194 %
Total Number of Instances            1282

```

```

==== Detailed Accuracy By Class ====
      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
      0.997    0.240    0.985     0.997    0.991     0.834    0.982     0.997    negative
      0.760    0.003    0.934     0.760    0.838     0.834    0.982     0.901    sick
Weighted Avg.  0.983    0.226    0.982     0.983    0.982     0.834    0.982     0.991

==== Confusion Matrix ====
      a      b  <-- classified as
1203    4 |    a = negative
      18   57 |    b = sick

```

## Leave one out fold

```

==== Run information ====
Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B40-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  258
              [list of attributes omitted]
Test mode:   3772-fold cross-validation

==== Classifier model (full training set) ====
PART decision list
-----
T3_4 = '(1.15-inf)' AND
FTI_38 = '(-inf-165.5)' AND
TT4_2 = '(62.5-inf)': negative (3118.35/1.0)

referral_source = other AND
age_35 = '(-inf-72.5)' AND
sick = f AND
TSH_measured = t AND
T4U_25 = '(1.025-inf)': negative (78.35)

T3_measured = f: negative (115.58/3.0)

T3_4 = '(1.15-inf)' AND
T3_19 = '(2.65-inf)': negative (58.57)

T3_4 = '(1.15-inf)' AND
TT4_36 = '(-inf-150.5)' AND
T4U_12 = '(-inf-0.895)': negative (44.38)

TT4_1 = '(-inf-46.5)' AND
age_38 = '(-inf-78.5]' AND
TSH_38 = '(12.55-inf)': negative (38.46/1.0)

TT4_39 = '(-inf-192.5)' AND
referral_source = SVI AND
TSH_39 = '(-inf-29.5)' AND
TSH_38 = '(-inf-12.55)' AND
FTI_37 = '(-inf-153.5)' AND
TT4_1 = '(46.5-inf)' AND
T4U_33 = '(-inf-1.135)' AND
TT4_2 = '(62.5-inf)': sick (141.18/0.25)

sick = f AND
TT4_39 = '(192.5-inf)': negative (8.05)

sick = f AND
age_1 = '(16.5-inf)' AND
T3_18 = '(-inf-2.55)' AND
T4U_measured = f: negative (7.72/0.9)

sick = t: sick (7.13/0.01)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
T3_18 = '(-inf-2.55)' AND
age_39 = '(82.5-inf)' AND
TT4_32 = '(-inf-132.5)': sick (6.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_30 = '(-inf-1.055)' AND
T3_18 = '(-inf-2.55)' AND
age_39 = '(82.5-inf)' AND
TT4_32 = '(-inf-132.5)': sick (6.0)

```

```

T4U_28 = '(-inf-1.055)' AND
T3_18 = '(-inf-2.55)' AND
referral_source = other AND
FTI_2 = '(68.5-inf)' AND
age_37 = '(-inf-76.5)' AND
TSH_34 = '(-inf-4.05)' AND
T3_1 = '(0.55-inf)' AND
T3_9 = '(-inf-1.65)' AND
age_33 = '(-inf-70.5)': negative (21.67/1.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
psych = f AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
query_hyperthyroid = f AND
I131_treatment = f AND
T3_18 = '(-inf-2.55)' AND
FTI_35 = '(-inf-141.5)' AND
T3_10 = '(-inf-1.75)' AND
TT4_5 = '(75.5-inf)': sick (12.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
psych = f AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
query_hyperthyroid = f AND
I131_treatment = f AND
T3_18 = '(-inf-2.55)' AND
T4U_13 = '(0.905-inf)' AND
T4U_22 = '(-inf-0.995)' AND
T3_5 = '(1.25-inf)': negative (15.0/1.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
psych = f AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_28 = '(-inf-1.055)' AND
query_hyperthyroid = f AND
I131_treatment = f AND
T3_18 = '(-inf-2.55)' AND
query_on_thyroxine = f AND
age_26 = '(-inf-60.5)' AND
age_8 = '(32.5-inf)' AND
T3_9 = '(-inf-1.65)' AND
TT4_10 = '(-inf-87.5)': sick (18.0/3.0)

age_1 = '(16.5-inf)' AND
T4U_34 = '(-inf-1.155)' AND
psych = f AND
T4U_32 = '(-inf-1.115)' AND
T4U_31 = '(-inf-1.095)' AND
T4U_27 = '(-inf-1.045)' AND
query_hyperthyroid = f AND
I131_treatment = f AND
T3_18 = '(-inf-2.55)' AND
query_on_thyroxine = f AND
T3_4 = '(1.15-inf)' AND
on_thyroxine = f AND
TSH_7 = '(-inf-0.205)': sick (11.0)

psych = f AND
query_on_thyroxine = f AND
T3_18 = '(-inf-2.55)' AND
TSH_4 = '(0.085-inf)' AND
query_hyperthyroid = f AND
TT4_21 = '(106.5-inf)': negative (13.02)

T4U_34 = '(-inf-1.155)' AND
TT4_37 = '(-inf-159.5)' AND
TSH_4 = '(0.085-inf)' AND
TT4_2 = '(62.5-inf)': sick (14.32/1.4)

TSH_37 = '(-inf-8.25)' AND
T4U_1 = '(0.685-inf)' AND
on_thyroxine = f: sick (17.14/2.05)

sex = F: negative (17.0/0.0)

TT4_3 = '(68.5-inf)': negative (4.72)

: sick (4.38/1.35)

```

```

Number of Rules : 22

Time taken to build model: 0.66 seconds

==== Stratified cross-validation ===
==== Summary ===

Correctly Classified Instances      3712      98.4093 %
Incorrectly Classified Instances   60       1.5907 %
Kappa statistic                   0.8588
Mean absolute error               0.0189
Root mean squared error          0.1225
Relative absolute error           16.413  %
Root relative squared error     51.0966 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall    F-Measure   MCC     ROC Area   PRC Area   Class
      0.993     0.152     0.990     0.993     0.992     0.859     0.915     0.986   negative
      0.848     0.007     0.887     0.848     0.867     0.859     0.915     0.827   sick
Weighted Avg.      0.984     0.143     0.984     0.984     0.984     0.859     0.915     0.976

==== Confusion Matrix ====

      a      b  <-- classified as
3516   25 |      a = negative
      35  196 |      b = sick

```

### 3.2.5 Conclusion

PART							
Preprocessing method	Cross Validation			Percentage Split			Leave One Out Fold
	5 folds	10 folds	15 folds	50.00%	60.00%	66.00%	
With Original Dataset	98.3033%	98.6214%	98.5949%	98.6214%	97.8131%	98.7520%	98.8865%
Equal frequency binning	40 bins	98.5419%	98.2768%	97.9586%	98.6214%	98.0782%	98.2839%
	10 bins	98.1442%	97.7466%	97.9056%	97.8261%	97.2830%	97.9719%
	5 bins	94.0880%	93.6108%	94.3001%	93.9555%	94.0358%	93.2917%
	2 bins	93.7699%	93.4517%	93.7434%	93.6373%	93.9032%	94.0718%
							93.6108%

PART was able to benefit from **discretizing with equal frequency binning and 40 bins** in two tests in Cross validation 5 folds and in Percentage split 60%. same as before what we can conclude is that increasing the number of bins had a positive impact on some of the cases however the price in terms of execution time is huge and with larger datasets this can consume time.

## 3.3 Supervised discretization

used filtered classifier from meta filter on the classifier tab, with Discretizing filter where Makebinary is set to true and PART as classifier.

### Cross Validation 5 folds

```

==== Run information ===

Scheme:      weka.classifiers.meta.FilteredClassifier -F "weka.filters.supervised.attribute.Discretize -D -R first-last -precision 6" -S 1 -W PART
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              T131_treatment

```

```

query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH
T3_measured
T3
TT4_measured
TT4
T4U_measured
T4U
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Test mode: 5-fold cross-validation

==== Classifier model (full training set) ====

FilteredClassifier using weka.classifiers.rules.PART -C 0.25 -M 2 on data filtered through weka.filters.supervised.attribute.Discretize -D -R

Filtered Header
@relation sick-weka.filters.supervised.attribute.Discretize-D-Rfirst-last-precision6

@attribute age_1 (''(-inf-43.5)'',''(43.5-inf)'')
@attribute age_2 (''(-inf-69.5)'',''(69.5-inf)'')
@attribute sex {F,M}
@attribute on_thyroxine {f,t}
@attribute query_on_thyroxine {f,t}
@attribute on_antithyroid_medication {f,t}
@attribute sick {f,t}
@attribute pregnant {f,t}
@attribute thyroid_surgery {f,t}
@attribute T131_treatment {f,t}
@attribute query_hypothyroid {f,t}
@attribute query_hyperthyroid {f,t}
@attribute lithium {f,t}
@attribute goitre {f,t}
@attribute tumor {f,t}
@attribute hypopituitary {f,t}
@attribute psych {f,t}
@attribute TSH_measured {t,f}
@attribute TSH {'All'}
@attribute T3_measured {t,f}
@attribute T3_1 (''(-inf-1.15)'',''(1.15-inf)'')
@attribute TT4_measured {t,f}
@attribute TT4_1 (''(-inf-87.5)'',''(87.5-inf)'')
@attribute T4U_measured {t,f}
@attribute T4U_1 (''(-inf-0.775)'',''(0.775-inf)'')
@attribute T4U_2 (''(-inf-0.885)'',''(0.885-inf)'')
@attribute T4U_3 (''(-inf-1.145)'',''(1.145-inf)'')
@attribute FTI_measured {t,f}
@attribute FTI {'All'}
@attribute TBG_measured {f}
@attribute TBG {'All'}
@attribute referral_source {SVHC,other,SVI,STMW,SVHD}
@attribute Class {negative,sick}

@data

Classifier Model
PART decision list
-----
T3_1 = '(1.15-inf)': negative (3381.36/24.69)

T3_measured = t AND
referral_source = SVI AND
sick = f AND
TT4_1 = '(87.5-inf)' AND
on_thyroxine = f: sick (85.0/8.0)

T3_measured = f: negative (79.64/0.31)

sick = f AND
T4U_measured = t AND
T4U_3 = '(-inf-1.145)' AND
referral_source = other AND
age_2 = '(-inf-69.5)' AND
on_thyroxine = f: negative (41.0/3.0)

```

```

T4U_3 = '(-inf-1.145]' AND
T4U_measured = t AND
sick = f AND
psych = f AND
T4U_2 = '(-inf-0.885]' AND
on_thyroxine = f AND
query_on_thyroxine = f AND
referral_source = SVI AND
query_hypothyroid = f: sick (74.0/16.0)

sick = t: sick (25.0)

T4U_measured = t AND
T4U_3 = '(-inf-1.145]' AND
age_1 = '(43.5-inf)' AND
psych = f AND
referral_source = other AND
on_thyroxine = f: sick (16.0/4.0)

psych = f AND
T4U_measured = t AND
T4U_3 = '(1.145-inf)': negative (10.0)

T4U_measured = f: negative (7.0)

age_1 = '(43.5-inf)' AND
psych = f AND
referral_source = SVHC: sick (6.0)

age_1 = '(-inf-43.5)': sick (4.0)

psych = t: sick (4.0)

referral_source = other AND
TT4_1 = '(87.5-inf)': negative (3.0)

referral_source = SVI AND
query_hypothyroid = f AND
T4U_2 = '(0.885-inf)': negative (14.0/2.0)

referral_source = SVI AND
age_2 = '(-inf-69.5)': sick (7.0)

T4U_2 = '(-inf-0.885)': negative (9.0/3.0)

: sick (6.0/1.0)

Number of Rules : 17

Time taken to build model: 0.36 seconds

==== Stratified cross-validation ====
==== Summary ====

    Correctly Classified Instances      3698          98.0382 %
    Incorrectly Classified Instances     74           1.9618 %
    Kappa statistic                      0.8287
    Mean absolute error                  0.0329
    Root mean squared error              0.134
    Relative absolute error              28.519 %
    Root relative squared error         55.9034 %
    Total Number of Instances            3772

==== Detailed Accuracy By Class ====

    TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
    0.990     0.165     0.989      0.990     0.990      0.829     0.924     0.989     negative
    0.835     0.010     0.843      0.835     0.839      0.829     0.924     0.765     sick
    Weighted Avg.   0.980     0.155     0.980      0.980     0.980      0.829     0.924     0.976

==== Confusion Matrix ====

    a     b  <-- classified as
3505   36 |     a = negative
    38   193 |     b = sick

```

## Cross Validation 10 folds

```

==== Run information ====
Scheme:      weka.classifiers.meta.FilteredClassifier -F "weka.filters.supervised.attribute.Discretize -D -R first-last -precision 6" -S 1 -W
Relation:    sick

```

```

Instances: 3772
Attributes: 30
age
sex
on_thyroxine
query_on_thyroxine
on_antithyroid_medication
sick
pregnant
thyroid_surgery
I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH
T3_measured
T3
TT4_measured
TT4
T4U_measured
T4U
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Test mode: 10-fold cross-validation

==== Classifier model (full training set) ====

FilteredClassifier using weka.classifiers.rules.PART -C 0.25 -M 2 on data filtered through weka.filters.supervised.attribute.Discretize -D -R ...

Filtered Header
@relation sick-weka.filters.supervised.attribute.Discretize-D-Rfirst-last-precision6

@attribute age_1 (''(-inf-43.5)'',''(43.5-inf)'')
@attribute age_2 (''(-inf-69.5)'',''(69.5-inf)'')
@attribute sex {F,M}
@attribute on_thyroxine {f,t}
@attribute query_on_thyroxine {f,t}
@attribute on_antithyroid_medication {f,t}
@attribute sick {f,t}
@attribute pregnant {f,t}
@attribute thyroid_surgery {f,t}
@attribute I131_treatment {f,t}
@attribute query_hypothyroid {f,t}
@attribute query_hyperthyroid {f,t}
@attribute lithium {f,t}
@attribute goitre {f,t}
@attribute tumor {f,t}
@attribute hypopituitary {f,t}
@attribute psych {f,t}
@attribute TSH_measured {t,f}
@attribute TSH {'All'}
@attribute T3_measured {t,f}
@attribute T3_1 (''(-inf-1.15)'',''(1.15-inf)'')
@attribute TT4_measured {t,f}
@attribute TT4_1 (''(-inf-87.5)'',''(87.5-inf)'')
@attribute T4U_measured {t,f}
@attribute T4U_1 (''(-inf-0.775)'',''(0.775-inf)'')
@attribute T4U_2 (''(-inf-0.885)'',''(0.885-inf)'')
@attribute T4U_3 (''(-inf-1.145)'',''(1.145-inf)'')
@attribute FTI_measured {t,f}
@attribute FTI {'All'}
@attribute TBG_measured {f}
@attribute TBG {'All'}
@attribute referral_source {SVHC,other,SVI,STMW,SVHD}
@attribute Class {negative,sick}

@data

Classifier Model
PART decision list
-----
T3_1 = '(1.15-inf)': negative (3381.36/24.69)

T3_measured = t AND
referral_source = SVI AND
sick = f AND

```

```

TT4_1 = '(87.5-inf)' AND
on_thyroxine = f: sick (85.0/8.0)

T3_measured = f: negative (79.64/0.31)

sick = f AND
T4U_measured = t AND
T4U_3 = '(-inf-1.145]' AND
referral_source = other AND
age_2 = '(-inf-69.5]' AND
on_thyroxine = f: negative (41.0/3.0)

T4U_3 = '(-inf-1.145]' AND
T4U_measured = t AND
sick = f AND
psych = f AND
T4U_2 = '(-inf-0.885]' AND
on_thyroxine = f AND
query_on_thyroxine = f AND
referral_source = SVI AND
query_hypothyroid = f: sick (74.0/16.0)

sick = t: sick (25.0)

T4U_measured = t AND
T4U_3 = '(-inf-1.145]' AND
age_1 = '(43.5-inf)' AND
psych = f AND
referral_source = other AND
on_thyroxine = f: sick (16.0/4.0)

psych = f AND
T4U_measured = t AND
T4U_3 = '(1.145-inf)': negative (10.0)

T4U_measured = f: negative (7.0)

age_1 = '(43.5-inf)' AND
psych = f AND
referral_source = SVHC: sick (6.0)

age_1 = '(-inf-43.5]': sick (4.0)

psych = t: sick (4.0)

referral_source = other AND
TT4_1 = '(87.5-inf)': negative (3.0)

referral_source = SVI AND
query_hypothyroid = f AND
T4U_2 = '(0.885-inf)': negative (14.0/2.0)

referral_source = SVI AND
age_2 = '(-inf-69.5]': sick (7.0)

T4U_2 = '(-inf-0.885]': negative (9.0/3.0)

: sick (6.0/1.0)

Number of Rules : 17

Time taken to build model: 0.14 seconds

==== Stratified cross-validation ===
==== Summary ===

    Correctly Classified Instances      3695          97.9586 %
    Incorrectly Classified Instances     77           2.0414 %
    Kappa statistic                      0.8206
    Mean absolute error                  0.0327
    Root mean squared error              0.1352
    Relative absolute error              28.4064 %
    Root relative squared error          56.4073 %
    Total Number of Instances            3772

==== Detailed Accuracy By Class ===

    TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
    0.990     0.177     0.988     0.990     0.989     0.821     0.909     0.985     negative
    0.823     0.010     0.841     0.823     0.832     0.821     0.909     0.765     sick
    Weighted Avg.   0.980     0.167     0.979     0.980     0.979     0.821     0.909     0.971

==== Confusion Matrix ====

    a      b  <-- classified as
3505   36 |      a = negative

```

```
41 190 | b = sick
```

## Cross Validation 15 folds

```
== Run information ==
Scheme:      weka.classifiers.meta.FilteredClassifier -F "weka.filters.supervised.attribute.Discretize -D -R first-last -precision 6" -S 1 -W
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   15-fold cross-validation

== Classifier model (full training set) ==
FilteredClassifier using weka.classifiers.rules.PART -C 0.25 -M 2 on data filtered through weka.filters.supervised.attribute.Discretize -D -R first-last-precision6

Filtered Header
@relation sick-weka.filters.supervised.attribute.Discretize-D-Rfirst-last-precision6

@attribute age_1 ('`(-inf-43.5)`','`(43.5-inf)`')
@attribute age_2 ('`(-inf-69.5)`','`(69.5-inf)`')
@attribute sex {F,M}
@attribute on_thyroxine {f,t}
@attribute query_on_thyroxine {f,t}
@attribute on_antithyroid_medication {f,t}
@attribute sick {f,t}
@attribute pregnant {f,t}
@attribute thyroid_surgery {f,t}
@attribute I131_treatment {f,t}
@attribute query_hypothyroid {f,t}
@attribute query_hyperthyroid {f,t}
@attribute lithium {f,t}
@attribute goitre {f,t}
@attribute tumor {f,t}
@attribute hypopituitary {f,t}
@attribute psych {f,t}
@attribute TSH_measured {t,f}
@attribute TSH {'`All`'}
@attribute T3_measured {t,f}
@attribute T3_1 ('`(-inf-1.15)`','`(1.15-inf)`')
@attribute TT4_measured {t,f}
@attribute TT4_1 ('`(-inf-87.5)`','`(87.5-inf)`')
@attribute T4U_measured {t,f}
@attribute T4U_1 ('`(-inf-0.775)`','`(0.775-inf)`')
@attribute T4U_2 ('`(-inf-0.885)`','`(0.885-inf)`')
@attribute T4U_3 ('`(-inf-1.145)`','`(1.145-inf)`')
@attribute FTI_measured {t,f}
@attribute FTI {'`All`'}
@attribute TBG_measured {f}
@attribute TBG {'`All`'}
@attribute referral_source {SVHC,other,SVI,STMW,SVHD}
@attribute Class {negative,sick}

@data
```

```

Classifier Model
PART decision list
-----
T3_1 = '(1.15-inf)': negative (3381.36/24.69)

T3_measured = t AND
referral_source = SVI AND
sick = f AND
TT4_1 = '(87.5-inf)' AND
on_thyroxine = f: sick (85.0/8.0)

T3_measured = f: negative (79.64/0.31)

sick = f AND
T4U_measured = t AND
T4U_3 = '(-inf-1.145]' AND
referral_source = other AND
age_2 = '(-inf-69.5]' AND
on_thyroxine = f: negative (41.0/3.0)

T4U_3 = '(-inf-1.145]' AND
T4U_measured = t AND
sick = f AND
psych = f AND
T4U_2 = '(-inf-0.885]' AND
on_thyroxine = f AND
query_on_thyroxine = f AND
referral_source = SVI AND
query_hypothyroid = f: sick (74.0/16.0)

sick = t: sick (25.0)

T4U_measured = t AND
T4U_3 = '(-inf-1.145]' AND
age_1 = '(43.5-inf)' AND
psych = f AND
referral_source = other AND
on_thyroxine = f: sick (16.0/4.0)

psych = f AND
T4U_measured = t AND
T4U_3 = '(1.145-inf)': negative (10.0)

T4U_measured = f: negative (7.0)

age_1 = '(43.5-inf)' AND
psych = f AND
referral_source = SVHC: sick (6.0)

age_1 = '(-inf-43.5]': sick (4.0)

psych = t: sick (4.0)

referral_source = other AND
TT4_1 = '(87.5-inf)': negative (3.0)

referral_source = SVI AND
query_hypothyroid = f AND
T4U_2 = '(0.885-inf)': negative (14.0/2.0)

referral_source = SVI AND
age_2 = '(-inf-69.5]': sick (7.0)

T4U_2 = '(-inf-0.885)': negative (9.0/3.0)

: sick (6.0/1.0)

Number of Rules : 17

Time taken to build model: 0.13 seconds

==== Stratified cross-validation ===
==== Summary ===

Correctly Classified Instances      3695          97.9586 %
Incorrectly Classified Instances    77           2.0414 %
Kappa statistic                   0.8206
Mean absolute error               0.0327
Root mean squared error           0.1357
Relative absolute error            28.3889 %
Root relative squared error       56.5865 %
Total Number of Instances         3772

```

```

==== Detailed Accuracy By Class ====
      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
      0.990    0.177    0.988    0.990    0.989    0.821    0.909    0.985    negative
      0.823    0.010    0.841    0.823    0.832    0.821    0.909    0.759    sick
Weighted Avg.  0.980    0.167    0.979    0.980    0.979    0.821    0.909    0.971

==== Confusion Matrix ====
      a     b  <-- classified as
3505  36 |  a = negative
      41 190 |  b = sick

```

## Percentage Split 50%

```

==== Run information ====
Scheme:      weka.classifiers.meta.FilteredClassifier -F "weka.filters.supervised.attribute.Discretize -D -R first-last -precision 6" -S 1 -W
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 50.0% train, remainder test

==== Classifier model (full training set) ====
FilteredClassifier using weka.classifiers.rules.PART -C 0.25 -M 2 on data filtered through weka.filters.supervised.attribute.Discretize -D -R first-last -precision 6
@relation sick-weka.filters.supervised.attribute.Discretize-D-Rfirst-last-precision6

@attribute age_1 ('`(-inf-43.5]`','`[43.5-inf)`')
@attribute age_2 ('`(-inf-69.5]`','`[69.5-inf)`')
@attribute sex {F,M}
@attribute on_thyroxine {f,t}
@attribute query_on_thyroxine {f,t}
@attribute on_antithyroid_medication {f,t}
@attribute sick {f,t}
@attribute pregnant {f,t}
@attribute thyroid_surgery {f,t}
@attribute I131_treatment {f,t}
@attribute query_hypothyroid {f,t}
@attribute query_hyperthyroid {f,t}
@attribute lithium {f,t}
@attribute goitre {f,t}
@attribute tumor {f,t}
@attribute hypopituitary {f,t}
@attribute psych {f,t}
@attribute TSH_measured {t,f}
@attribute TSH {'`All`'}
@attribute T3_measured {t,f}
@attribute T3_1 ('`(-inf-1.15]`','`[1.15-inf)`')
@attribute TT4_measured {t,f}
@attribute TT4_1 ('`(-inf-87.5]`','`[87.5-inf)`')
@attribute T4U_measured {t,f}

```

```

@attribute T4U_1 {'\(-inf-0.775]\', '\(0.775-inf]\'}
@attribute T4U_2 {'\(-inf-0.885]\', '\(0.885-inf]\'}
@attribute T4U_3 {'\(-inf-1.145]\', '\(1.145-inf]\'}
@attribute FTI_measured {t,f}
@attribute FTI {'\All\'}
@attribute TBG_measured {f}
@attribute TBG {'\All\'}
@attribute referral_source {SVHC,other,SVI,STMW,SVHD}
@attribute Class {negative,sick}

@data

Classifier Model
PART decision list
-----
T3_1 = '(1.15-inf)': negative (3381.36/24.69)

T3_measured = t AND
referral_source = SVI AND
sick = f AND
TT4_1 = '(87.5-inf)' AND
on_thyroxine = f: sick (85.0/8.0)

T3_measured = f: negative (79.64/0.31)

sick = f AND
T4U_measured = t AND
T4U_3 = '\(-inf-1.145]' AND
referral_source = other AND
age_2 = '\(-inf-69.5]' AND
on_thyroxine = f: negative (41.0/3.0)

T4U_3 = '\(-inf-1.145]' AND
T4U_measured = t AND
sick = f AND
psych = f AND
T4U_2 = '\(-inf-0.885]' AND
on_thyroxine = f AND
query_on_thyroxine = f AND
referral_source = SVI AND
query_hypothyroid = f: sick (74.0/16.0)

sick = t: sick (25.0)

T4U_measured = t AND
T4U_3 = '\(-inf-1.145]' AND
age_1 = '(43.5-inf)' AND
psych = f AND
referral_source = other AND
on_thyroxine = f: sick (16.0/4.0)

psych = f AND
T4U_measured = t AND
T4U_3 = '(1.145-inf)': negative (10.0)

T4U_measured = f: negative (7.0)

age_1 = '(43.5-inf)' AND
psych = f AND
referral_source = SVHC: sick (6.0)

age_1 = '\(-inf-43.5]': sick (4.0)

psych = t: sick (4.0)

referral_source = other AND
TT4_1 = '(87.5-inf)': negative (3.0)

referral_source = SVI AND
query_hypothyroid = f AND
T4U_2 = '(0.885-inf)': negative (14.0/2.0)

referral_source = SVI AND
age_2 = '\(-inf-69.5]': sick (7.0)

T4U_2 = '\(-inf-0.885]': negative (9.0/3.0)

: sick (6.0/1.0)

Number of Rules : 17

Time taken to build model: 0.13 seconds
==== Evaluation on test split ====

```

```

Time taken to test model on test split: 0.02 seconds

==== Summary ====

Correctly Classified Instances      1841          97.614 %
Incorrectly Classified Instances   45           2.386 %
Kappa statistic                   0.7838
Mean absolute error               0.0344
Root mean squared error          0.14
Relative absolute error           29.8173 %
Root relative squared error     58.0274 %
Total Number of Instances        1886

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC    ROC Area   PRC Area   Class
      0.991     0.248     0.984     0.991     0.987     0.785     0.902     0.986   negative
      0.752     0.009     0.846     0.752     0.796     0.785     0.902     0.732   sick
Weighted Avg.      0.976     0.233     0.975     0.976     0.975     0.785     0.902     0.970

==== Confusion Matrix ====

      a     b  <-- classified as
1753  16 |     a = negative
  29  88 |     b = sick

```

## Percentage Split 60%

```

==== Run information ====

Scheme:      weka.classifiers.meta.FilteredClassifier -F "weka.filters.supervised.attribute.Discretize -D -R first-last -precision 6" -S 1 -W
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 60.0% train, remainder test

==== Classifier model (full training set) ====

FilteredClassifier using weka.classifiers.rules.PART -C 0.25 -M 2 on data filtered through weka.filters.supervised.attribute.Discretize -D -R

Filtered Header
@relation sick-weka.filters.supervised.attribute.Discretize-D-Rfirst-last-precision6

@attribute age_1 ('`(-inf-43.5]`','`[43.5-inf)`')
@attribute age_2 ('`(-inf-69.5]`','`[69.5-inf)`')
@attribute sex {F,M}
@attribute on_thyroxine {f,t}
@attribute query_on_thyroxine {f,t}
@attribute on_antithyroid_medication {f,t}
@attribute sick {f,t}
@attribute pregnant {f,t}
@attribute thyroid_surgery {f,t}
@attribute I131_treatment {f,t}

```

```

@attribute query_hypothyroid {f,t}
@attribute query_hyperthyroid {f,t}
@attribute lithium {f,t}
@attribute goitre {f,t}
@attribute tumor {f,t}
@attribute hypopituitary {f,t}
@attribute psych {f,t}
@attribute TSH_measured {t,f}
@attribute TSH {'\All\''}
@attribute T3_measured {t,f}
@attribute T3_1 {'\`(-inf-1.15)\``,\``(1.15-inf)\``'}
@attribute TT4_measured {t,f}
@attribute TT4_1 {'\`(-inf-87.5]\``,\``(87.5-inf)\``'}
@attribute T4U_measured {t,f}
@attribute T4U_1 {'\`(-inf-0.775]\``,\``(0.775-inf)\``'}
@attribute T4U_2 {'\`(-inf-0.885]\``,\``(0.885-inf)\``'}
@attribute T4U_3 {'\`(-inf-1.145]\``,\``(1.145-inf)\``'}
@attribute FTI_measured {t,f}
@attribute FTI {'\All\''}
@attribute TBG_measured {f}
@attribute TBG {'\All\''}
@attribute referral_source {SVHC,other,SVI,STMW,SVHD}
@attribute Class {negative,sick}

@data

Classifier Model
PART decision list
-----
T3_1 = '(1.15-inf)': negative (3381.36/24.69)

T3_measured = t AND
referral_source = SVI AND
sick = f AND
TT4_1 = '(87.5-inf)' AND
on_thyroxine = f: sick (85.0/8.0)

T3_measured = f: negative (79.64/0.31)

sick = f AND
T4U_measured = t AND
T4U_3 = '(-inf-1.145]' AND
referral_source = other AND
age_2 = '(-inf-69.5]' AND
on_thyroxine = f: negative (41.0/3.0)

T4U_3 = '(-inf-1.145]' AND
T4U_measured = t AND
sick = f AND
psych = f AND
T4U_2 = '(-inf-0.885]' AND
on_thyroxine = f AND
query_on_thyroxine = f AND
referral_source = SVI AND
query_hypothyroid = f: sick (74.0/16.0)

sick = t: sick (25.0)

T4U_measured = t AND
T4U_3 = '(-inf-1.145]' AND
age_1 = '(43.5-inf)' AND
psych = f AND
referral_source = other AND
on_thyroxine = f: sick (16.0/4.0)

psych = f AND
T4U_measured = t AND
T4U_3 = '(1.145-inf)': negative (10.0)

T4U_measured = f: negative (7.0)

age_1 = '(43.5-inf)' AND
psych = f AND
referral_source = SVHC: sick (6.0)

age_1 = '(-inf-43.5)': sick (4.0)

psych = t: sick (4.0)

referral_source = other AND
TT4_1 = '(87.5-inf)': negative (3.0)

referral_source = SVI AND
query_hypothyroid = f AND
T4U_2 = '(0.885-inf)': negative (14.0/2.0)

```

```

referral_source = SVI AND
age_2 = '(-inf-69.5]': sick (7.0)

T4U_2 = '(-inf-0.885]': negative (9.0/3.0)
: sick (6.0/1.0)

Number of Rules : 17

Time taken to build model: 0.13 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0.04 seconds

==== Summary ===

Correctly Classified Instances      1478          97.9457 %
Incorrectly Classified Instances   31           2.0543 %
Kappa statistic                   0.8099
Mean absolute error               0.0333
Root mean squared error          0.1344
Relative absolute error          29.0362 %
Root relative squared error     57.0484 %
Total Number of Instances        1509

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC    ROC Area   PRC Area   Class
      0.991     0.202     0.987     0.991     0.989     0.810     0.941     0.993   negative
      0.798     0.009     0.845     0.798     0.821     0.810     0.941     0.798   sick
Weighted Avg.       0.979     0.191     0.979     0.979     0.979     0.810     0.941     0.981

==== Confusion Matrix ===

      a     b  <-- classified as
1407  13 |   a = negative
      18  71 |   b = sick

```

## Percentage Split 66%

```

==== Run information ===

Scheme:      weka.classifiers.meta.FilteredClassifier -F "weka.filters.supervised.attribute.Discretize -D -R first-last -precision 6" -S 1 -W
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              t131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 66.0% train, remainder test

==== Classifier model (full training set) ===

FilteredClassifier using weka.classifiers.rules.PART -C 0.25 -M 2 on data filtered through weka.filters.supervised.attribute.Discretize -D -R

```

```

Filtered Header
@relation sick-weka.filters.supervised.attribute.Discretize-D-Rfirst-last-precision6

@attribute age_1 {'\*(-inf-43.5]\', '\*(43.5-inf)\'}
@attribute age_2 {'\*(-inf-69.5]\', '\*(69.5-inf)\'}
@attribute sex {F,M}
@attribute on_thyroxine {f,t}
@attribute query_on_thyroxine {f,t}
@attribute on_antithyroid_medication {f,t}
@attribute sick {f,t}
@attribute pregnant {f,t}
@attribute thyroid_surgery {f,t}
@attribute I131_treatment {f,t}
@attribute query_hypothyroid {f,t}
@attribute query_hyperthyroid {f,t}
@attribute lithium {f,t}
@attribute goitre {f,t}
@attribute tumor {f,t}
@attribute hypopituitary {f,t}
@attribute psych {f,t}
@attribute TSH_measured {t,f}
@attribute TSH {'\*All\'}
@attribute T3_measured {t,f}
@attribute T3_1 {'\*(-inf-1.15]\', '\*(1.15-inf)\'}
@attribute TT4_measured {t,f}
@attribute TT4_1 {'\*(-inf-87.5]\', '\*(87.5-inf)\'}
@attribute T4U_measured {t,f}
@attribute T4U_1 {'\*(-inf-0.775]\', '\*(0.775-inf)\'}
@attribute T4U_2 {'\*(-inf-0.885]\', '\*(0.885-inf)\'}
@attribute T4U_3 {'\*(-inf-1.145]\', '\*(1.145-inf)\'}
@attribute FTI_measured {t,f}
@attribute FTI {'\*All\'}
@attribute TBG_measured {f}
@attribute TBG {'\*All\'}
@attribute referral_source {SVHC,other,SVI,STMW,SVHD}
@attribute Class {negative,sick}

@data

Classifier Model
PART decision list
-----
T3_1 = '(1.15-inf)': negative (3381.36/24.69)

T3_measured = t AND
referral_source = SVI AND
sick = f AND
TT4_1 = '(87.5-inf)' AND
on_thyroxine = f: sick (85.0/8.0)

T3_measured = f: negative (79.64/0.31)

sick = f AND
T4U_measured = t AND
T4U_3 = '(-inf-1.145]' AND
referral_source = other AND
age_2 = '(-inf-69.5]' AND
on_thyroxine = f: negative (41.0/3.0)

T4U_3 = '(-inf-1.145]' AND
T4U_measured = t AND
sick = f AND
psych = f AND
T4U_2 = '(-inf-0.885]' AND
on_thyroxine = f AND
query_on_thyroxine = f AND
referral_source = SVI AND
query_hypothyroid = f: sick (74.0/16.0)

sick = t: sick (25.0)

T4U_measured = t AND
T4U_3 = '(-inf-1.145]' AND
age_1 = '(43.5-inf)' AND
psych = f AND
referral_source = other AND
on_thyroxine = f: sick (16.0/4.0)

psych = f AND
T4U_measured = t AND
T4U_3 = '(1.145-inf)': negative (10.0)

T4U_measured = f: negative (7.0)

```

```

age_1 = '(43.5-inf)' AND
psych = f AND
referral_source = SVHC: sick (6.0)

age_1 = '(-inf-43.5]': sick (4.0)

psych = t: sick (4.0)

referral_source = other AND
TT4_1 = '(87.5-inf)': negative (3.0)

referral_source = SVI AND
query_hypothyroid = f AND
T4U_2 = '(0.885-inf)': negative (14.0/2.0)

referral_source = SVI AND
age_2 = '(-inf-69.5]': sick (7.0)

T4U_2 = '(-inf-0.885)': negative (9.0/3.0)

: sick (6.0/1.0)

Number of Rules : 17

Time taken to build model: 0.28 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0.04 seconds

==== Summary ===

Correctly Classified Instances      1253          97.7379 %
Incorrectly Classified Instances   29           2.2621 %
Kappa statistic                   0.7795
Mean absolute error               0.0351
Root mean squared error          0.1386
Relative absolute error          30.7574 %
Root relative squared error     59.0567 %
Total Number of Instances        1282

==== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC    ROC Area  PRC Area  Class
0.993     0.267    0.984     0.993    0.988     0.782    0.937     0.992  negative
0.733     0.007    0.859     0.733    0.791     0.782    0.937     0.758   sick
Weighted Avg.  0.977     0.252    0.976     0.977    0.977     0.782    0.937     0.979

==== Confusion Matrix ===

      a      b  <-- classified as
1198    9 |  a = negative
  20    55 |  b = sick

```

## Leave one out fold

```

==== Run information ===

Scheme:      weka.classifiers.meta.FilteredClassifier -F "weka.filters.supervised.attribute.Discretize -D -R first-last -precision 6" -S 1 -W
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured

```

```

TT4
T4U_measured
T4U
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Test mode: 3772-fold cross-validation

==== Classifier model (full training set) ===

FilteredClassifier using weka.classifiers.rules.PART -C 0.25 -M 2 on data filtered through weka.filters.supervised.attribute.Discretize -D -R ...

Filtered Header
@relation sick-weka.filters.supervised.attribute.Discretize-D-Rfirst-last-precision6

@attribute age_1 ('`(-inf-43.5]`','`[43.5-inf)`')
@attribute age_2 ('`(-inf-69.5]`','`[69.5-inf)`')
@attribute sex {F,M}
@attribute on_thyroxine {f,t}
@attribute query_on_thyroxine {f,t}
@attribute on_antithyroid_medication {f,t}
@attribute sick {f,t}
@attribute pregnant {f,t}
@attribute thyroid_surgery {f,t}
@attribute I131_treatment {f,t}
@attribute query_hypothyroid {f,t}
@attribute query_hyperthyroid {f,t}
@attribute lithium {f,t}
@attribute goitre {f,t}
@attribute tumor {f,t}
@attribute hypopituitary {f,t}
@attribute psych {f,t}
@attribute TSH_measured {t,f}
@attribute TSH {'`All`'}
@attribute T3_measured {t,f}
@attribute T3_1 ('`(-inf-1.15]`','`[1.15-inf)`')
@attribute TT4_measured {t,f}
@attribute TT4_1 ('`(-inf-87.5]`','`[87.5-inf)`')
@attribute T4U_measured {t,f}
@attribute T4U_1 ('`(-inf-0.775]`','`[0.775-inf)`')
@attribute T4U_2 ('`(-inf-0.885]`','`[0.885-inf)`')
@attribute T4U_3 ('`(-inf-1.145]`','`[1.145-inf)`')
@attribute FTI_measured {t,f}
@attribute FTI {'`All`'}
@attribute TBG_measured {f}
@attribute TBG {'`All`'}
@attribute referral_source {SVHC,other,SVI,STMW,SVHD}
@attribute Class {negative,sick}

@data

Classifier Model
PART decision list
-----
T3_1 = '(1.15-inf)': negative (3381.36/24.69)

T3_measured = t AND
referral_source = SVI AND
sick = f AND
TT4_1 = '(87.5-inf)' AND
on_thyroxine = f: sick (85.0/8.0)

T3_measured = f: negative (79.64/0.31)

sick = f AND
T4U_measured = t AND
T4U_3 = '(-inf-1.145]' AND
referral_source = other AND
age_2 = '(-inf-69.5]' AND
on_thyroxine = f: negative (41.0/3.0)

T4U_3 = '(-inf-1.145]' AND
T4U_measured = t AND
sick = f AND
psych = f AND
T4U_2 = '(-inf-0.885]' AND
on_thyroxine = f AND
query_on_thyroxine = f AND
referral_source = SVI AND
query_hypothyroid = f: sick (74.0/16.0)

sick = t: sick (25.0)

```

```

T4U_measured = t AND
T4U_3 = '(-inf-1.145]' AND
age_1 = '(43.5-inf)' AND
psych = f AND
referral_source = other AND
on_thyroxine = f: sick (16.0/4.0)

psych = f AND
T4U_measured = t AND
T4U_3 = '(1.145-inf)': negative (10.0)

T4U_measured = f: negative (7.0)

age_1 = '(43.5-inf)' AND
psych = f AND
referral_source = SVHC: sick (6.0)

age_1 = '(-inf-43.5)': sick (4.0)

psych = t: sick (4.0)

referral_source = other AND
TT4_1 = '(87.5-inf)': negative (3.0)

referral_source = SVI AND
query_hypothyroid = f AND
T4U_2 = '(0.885-inf)': negative (14.0/2.0)

referral_source = SVI AND
age_2 = '(-inf-69.5)': sick (7.0)

T4U_2 = '(-inf-0.885)': negative (9.0/3.0)

: sick (6.0/1.0)

Number of Rules : 17

Time taken to build model: 0.16 seconds

==== Stratified cross-validation ====
==== Summary ====

    Correctly Classified Instances      3697           98.0117 %
    Incorrectly Classified Instances     75            1.9883 %
    Kappa statistic                      0.8267
    Mean absolute error                  0.0326
    Root mean squared error              0.1338
    Relative absolute error              28.3164 %
    Root relative squared error          55.7719 %
    Total Number of Instances            3772

==== Detailed Accuracy By Class ====

    TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
    0.990     0.165     0.989     0.990     0.989     0.827     0.891     0.984     negative
    0.835     0.010     0.839     0.835     0.837     0.827     0.891     0.777     sick
    Weighted Avg.   0.980     0.155     0.980     0.980     0.980     0.827     0.891     0.971

==== Confusion Matrix ====

    a      b  <-- classified as
3504   37 |   a = negative
    38  193 |   b = sick

```

### 3.3.1 Conclusion

Preprocessing method	PART					
	Cross Validation			Percentage Split		
	5 folds	10 folds	15 folds	50.00%	60.00%	66.00%
With Original Dataset	98.3033%	98.6214%	98.5949%	98.6214%	97.8131%	98.7520%
Supervised	98.0382%	97.9586%	97.9586%	97.6140%	97.9457%	97.7379%
						98.0117%

Supervised Desensitizing didn't help much and it was able to beat the original dataset in a single test only Percentage split 60%.

## 3.4 Attribute Selection

### 3.4.1 Using Filters

From the Select attribute tab we choose the **GainRatioAttributeEval** as the attribute evaluator Ranker as the search method with 10 best attributes will be used in the tests. rest of options are left to default settings.

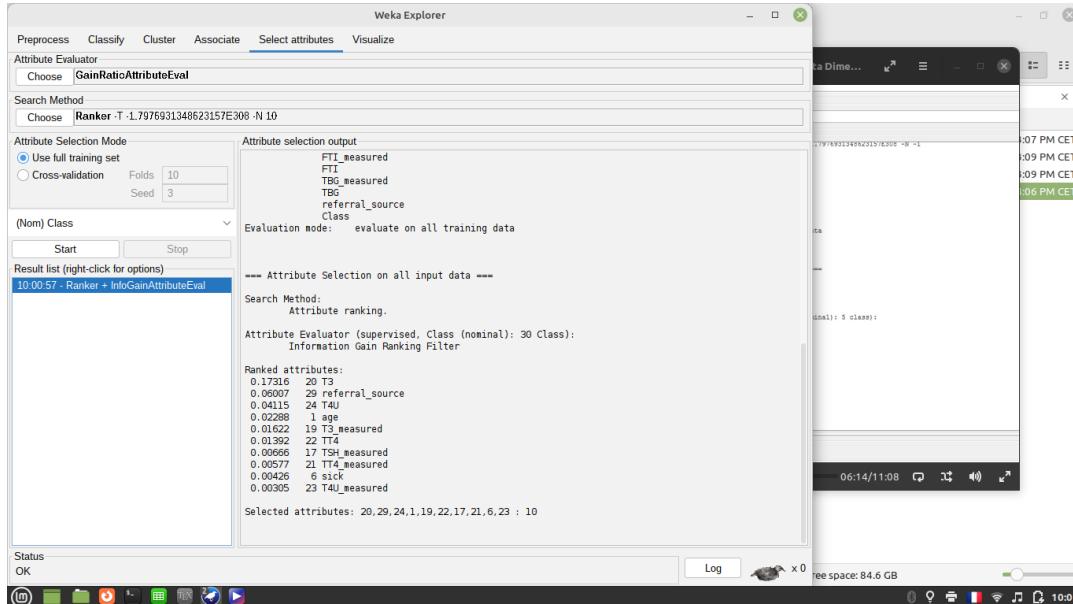


FIGURE 3.1: PART Filters1

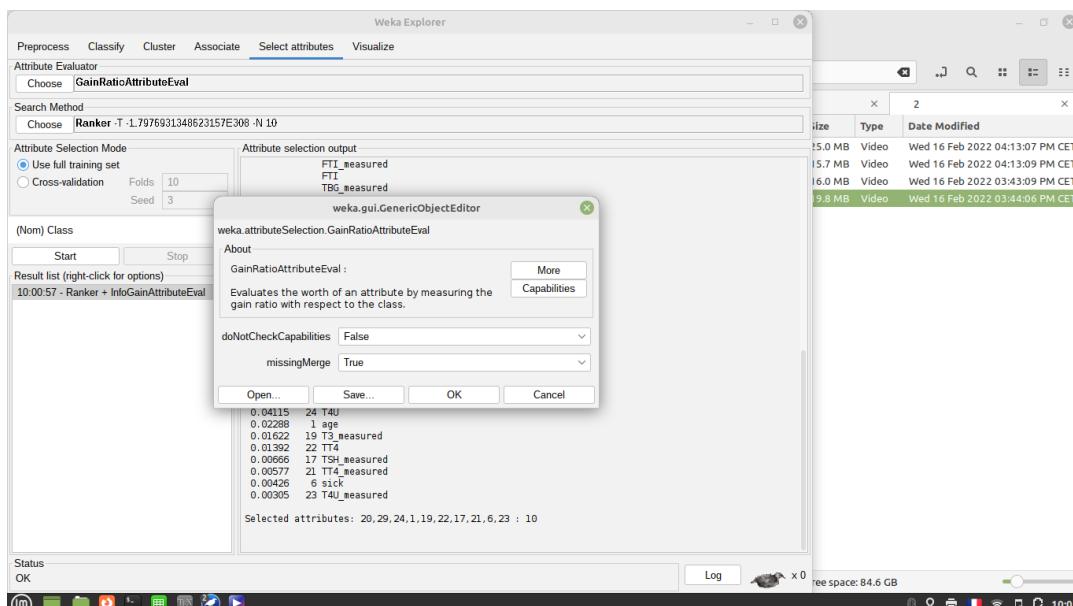


FIGURE 3.2: PART Filters2

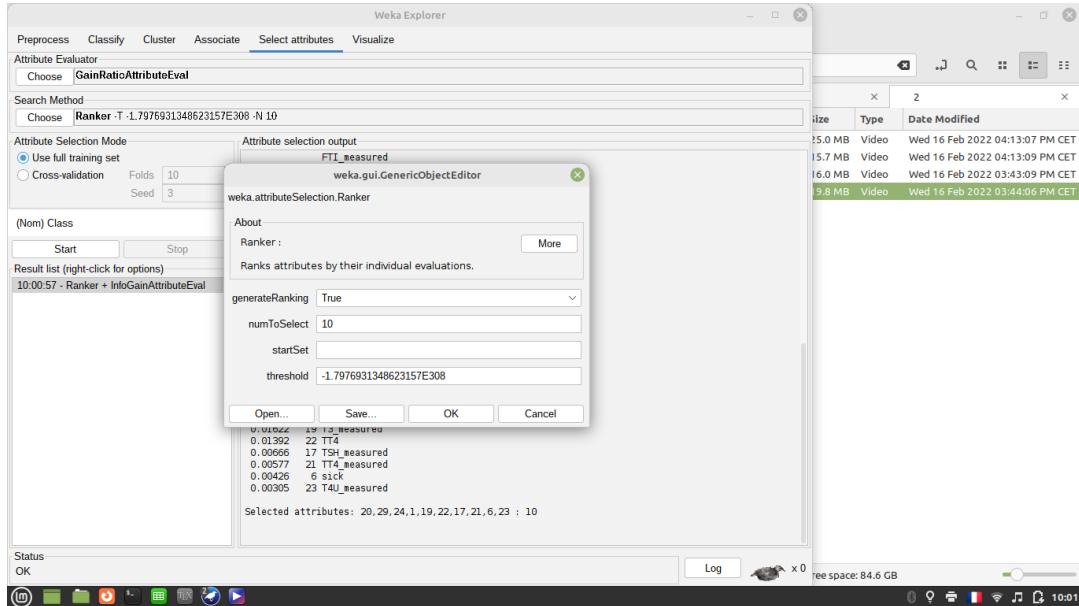


FIGURE 3.3: PART Filters3

```
==== Run information ====
Evaluator: weka.attributeSelection.InfoGainAttributeEval
Search: weka.attributeSelection.Ranker -T -1.7976931348623157E308 -N 10
Relation: sick
Instances: 3772
Attributes: 30
age
sex
on_thyroxine
query_on_thyroxine
on_antithyroid_medication
sick
pregnant
thyroid_surgery
I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH
T3_measured
T3
TT4_measured
TT4
T4U_measured
T4U
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Evaluation mode: evaluate on all training data

==== Attribute Selection on all input data ====
Search Method:
Attribute ranking.

Attribute Evaluator (supervised, Class (nominal): 30 Class):
Information Gain Ranking Filter

Ranked attributes:
0.17316 20 T3
0.06007 29 referral_source
0.04115 24 T4U
```

```

0.02288    1 age
0.01622    19 T3_measured
0.01392    22 TT4
0.00666    17 TSH_measured
0.00577    21 TT4_measured
0.00426    6 sick
0.00305    23 T4U_measured

Selected attributes: 20,29,24,1,19,22,17,21,6,23 : 10

```

## Cross Validation 5 folds

```

==== Run information ====
Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Remove-R2-5,7-16,18,25-28
Instances:   3772
Attributes:  11
              age
              sick
              TSH_measured
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              referral_source
              Class
Test mode:   5-fold cross-validation

==== Classifier model (full training set) ====
PART decision list
-----
T3 > 1.1 AND
TT4 <= 151 AND
TT4 > 56: negative (2986.67)

referral_source = other AND
age <= 72 AND
T4U > 1.02: negative (203.82)

TT4 > 145 AND
T3 > 2.6: negative (117.75)

T3_measured = f AND
referral_source = other: negative (60.45/1.0)

TT4 > 53 AND
TT4 <= 136 AND
T3_measured = t AND
referral_source = SVI: sick (168.11/6.11)

sick = f AND
age <= 70 AND
T4U_measured = t AND
referral_source = other: negative (71.0/15.0)

age <= 32: negative (22.4)

sick = f AND
T4U > 1.12: negative (15.91/0.23)

sick = f AND
referral_source = SVHC: sick (10.04/2.04)

sick = f AND
T4U_measured = f: negative (11.27/1.77)

referral_source = SVI AND
T3_measured = f AND
TT4 > 65: negative (4.94)

sick = f AND
referral_source = SVI AND
T3_measured = t: negative (63.05/20.0)

TT4 <= 137 AND
sick = f AND
referral_source = other AND
TT4 > 62: sick (13.0/1.0)

```

```

TT4 > 131: negative (10.48/1.0)
sick = t: sick (5.0)
referral_source = other: negative (4.0/1.0)
: sick (4.1/0.1)

Number of Rules : 17

Time taken to build model: 0.07 seconds

==== Stratified cross-validation ====
==== Summary ===

Correctly Classified Instances      3704          98.1972 %
Incorrectly Classified Instances    68           1.8028 %
Kappa statistic                   0.8345
Mean absolute error               0.0242
Root mean squared error          0.1246
Relative absolute error           20.9787 %
Root relative squared error     51.9664 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC     ROC Area   PRC Area   Class
      0.994     0.203     0.987     0.994     0.990     0.836     0.980     0.998   negative
      0.797     0.006     0.898     0.797     0.844     0.836     0.980     0.821   sick
Weighted Avg.      0.982     0.191     0.981     0.982     0.981     0.836     0.980     0.987

==== Confusion Matrix ====

      a     b  <-- classified as
3520   21 |     a = negative
      47  184 |     b = sick

```

## Cross Validation 10 folds

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Remove-R2-5,7-16,18,25-28
Instances:   3772
Attributes:  11
              age
              sick
              TSH_measured
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              referral_source
              Class
Test mode:   10-fold cross-validation

==== Classifier model (full training set) ===

PART decision list
-----
T3 > 1.1 AND
TT4 <= 151 AND
TT4 > 56: negative (2986.67)

referral_source = other AND
age <= 72 AND
T4U > 1.02: negative (203.82)

TT4 > 145 AND
T3 > 2.6: negative (117.75)

T3_measured = f AND
referral_source = other: negative (60.45/1.0)

TT4 > 53 AND
TT4 <= 136 AND
T3_measured = t AND
referral_source = SVI: sick (168.11/6.11)

sick = f AND

```

```

age <= 70 AND
T4U_measured = t AND
referral_source = other: negative (71.0/15.0)

age <= 32: negative (22.4)

sick = f AND
T4U > 1.12: negative (15.91/0.23)

sick = f AND
referral_source = SVHC: sick (10.04/2.04)

sick = f AND
T4U_measured = f: negative (11.27/1.77)

referral_source = SVI AND
T3_measured = f AND
TT4 > 65: negative (4.94)

sick = f AND
referral_source = SVI AND
T3_measured = t: negative (63.05/20.0)

TT4 <= 137 AND
sick = f AND
referral_source = other AND
TT4 > 62: sick (13.0/1.0)

TT4 > 131: negative (10.48/1.0)

sick = t: sick (5.0)

referral_source = other: negative (4.0/1.0)

: sick (4.1/0.1)

Number of Rules : 17

Time taken to build model: 0.05 seconds

==== Stratified cross-validation ====
==== Summary ====

    Correctly Classified Instances      3705           98.2238 %
    Incorrectly Classified Instances   67            1.7762 %
    Kappa statistic                   0.838
    Mean absolute error              0.0245
    Root mean squared error          0.1215
    Relative absolute error          21.226 %
    Root relative squared error     50.6782 %
    Total Number of Instances        3772

==== Detailed Accuracy By Class ====

    TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
    0.994     0.195     0.987     0.994     0.991     0.839     0.984     0.998     negative
    0.805     0.006     0.894     0.805     0.847     0.839     0.984     0.864     sick
    Weighted Avg.   0.982     0.183     0.982     0.982     0.982     0.839     0.984     0.990

==== Confusion Matrix ====

    a      b  <- classified as
3519   22 |  a = negative
    45   186 |  b = sick

```

## Cross Validation 15 folds

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Remove-R2-5,7-16,18,25-28
Instances:   3772
Attributes:  11
             age
             sick
             TSH_measured
             T3_measured
             T3
             TT4_measured
             TT4
             T4U_measured
             T4U
             referral_source

```

```

        Class
Test mode: 15-fold cross-validation

==== Classifier model (full training set) ====

PART decision list
-----

T3 > 1.1 AND
TT4 <= 151 AND
TT4 > 56: negative (2986.67)

referral_source = other AND
age <= 72 AND
T4U > 1.02: negative (203.82)

TT4 > 145 AND
T3 > 2.6: negative (117.75)

T3_measured = f AND
referral_source = other: negative (60.45/1.0)

TT4 > 53 AND
TT4 <= 136 AND
T3_measured = t AND
referral_source = SVI: sick (168.11/6.11)

sick = f AND
age <= 70 AND
T4U_measured = t AND
referral_source = other: negative (71.0/15.0)

age <= 32: negative (22.4)

sick = f AND
T4U > 1.12: negative (15.91/0.23)

sick = f AND
referral_source = SVHC: sick (10.04/2.04)

sick = f AND
T4U_measured = f: negative (11.27/1.77)

referral_source = SVI AND
T3_measured = f AND
TT4 > 65: negative (4.94)

sick = f AND
referral_source = SVI AND
T3_measured = t: negative (63.05/20.0)

TT4 <= 137 AND
sick = f AND
referral_source = other AND
TT4 > 62: sick (13.0/1.0)

TT4 > 131: negative (10.48/1.0)

sick = t: sick (5.0)

referral_source = other: negative (4.0/1.0)

: sick (4.1/0.1)

Number of Rules : 17

Time taken to build model: 0.05 seconds

==== Stratified cross-validation ====
==== Summary ====

    Correctly Classified Instances      3697           98.0117 %
    Incorrectly Classified Instances     75            1.9883 %
    Kappa statistic                      0.8155
    Mean absolute error                  0.0266
    Root mean squared error              0.1282
    Relative absolute error              23.1091 %
    Root relative squared error          53.4755 %
    Total Number of Instances            3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.994     0.229     0.985     0.994     0.989     0.818     0.972     0.997     negative
      0.771     0.006     0.890     0.771     0.826     0.818     0.972     0.816     sick
Weighted Avg.      0.980     0.216     0.979     0.980     0.979     0.818     0.972     0.986

```

```
==== Confusion Matrix ====
      a     b  <-- classified as
3519   22 |    a = negative
      53  178 |    b = sick
```

## Percentage Split 50%

```
==== Run information ====
Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Remove-R2-5,7-16,18,25-28
Instances:   3772
Attributes:  11
              age
              sick
              TSH_measured
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              referral_source
              Class
Test mode:   split 50.0% train, remainder test

==== Classifier model (full training set) ====
PART decision list
-----
T3 > 1.1 AND
TT4 <= 151 AND
TT4 > 56: negative (2986.67)

referral_source = other AND
age <= 72 AND
T4U > 1.02: negative (203.82)

TT4 > 145 AND
T3 > 2.6: negative (117.75)

T3_measured = f AND
referral_source = other: negative (60.45/1.0)

TT4 > 53 AND
TT4 <= 136 AND
T3_measured = t AND
referral_source = SVI: sick (168.11/6.11)

sick = f AND
age <= 70 AND
T4U_measured = t AND
referral_source = other: negative (71.0/15.0)

age <= 32: negative (22.4)

sick = f AND
T4U > 1.12: negative (15.91/0.23)

sick = f AND
referral_source = SVHC: sick (10.04/2.04)

sick = f AND
T4U_measured = f: negative (11.27/1.77)

referral_source = SVI AND
T3_measured = f AND
TT4 > 65: negative (4.94)

sick = f AND
referral_source = SVI AND
T3_measured = t: negative (63.05/20.0)

TT4 <= 137 AND
sick = f AND
referral_source = other AND
TT4 > 62: sick (13.0/1.0)

TT4 > 131: negative (10.48/1.0)

sick = t: sick (5.0)
```

```

referral_source = other: negative (4.0/1.0)
: sick (4.1/0.1)

Number of Rules : 17

Time taken to build model: 0.05 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

==== Summary ===

Correctly Classified Instances      1850          98.0912 %
Incorrectly Classified Instances   36           1.9088 %
Kappa statistic                   0.832
Mean absolute error               0.0242
Root mean squared error          0.1309
Relative absolute error          20.9661 %
Root relative squared error     54.2619 %
Total Number of Instances        1886

==== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC    ROC Area  PRC Area  Class
          0.992    0.179    0.988    0.992    0.990    0.832    0.964    0.996  negative
          0.821    0.008    0.865    0.821    0.842    0.832    0.965    0.779    sick
Weighted Avg.       0.981    0.169    0.981    0.981    0.981    0.832    0.964    0.982

==== Confusion Matrix ===

      a     b  <-- classified as
1754  15 |     a = negative
      21  96 |     b = sick

```

## Percentage Split 60%

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Remove-R2-5,7-16,18,25-28
Instances:   3772
Attributes:  11
              age
              sick
              TSH_measured
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              referral_source
              Class
Test mode:   split 60.0% train, remainder test

==== Classifier model (full training set) ===

PART decision list
-----
T3 > 1.1 AND
TT4 <= 151 AND
TT4 > 56: negative (2986.67)

referral_source = other AND
age <= 72 AND
T4U > 1.02: negative (203.82)

TT4 > 145 AND
T3 > 2.6: negative (117.75)

T3_measured = f AND
referral_source = other: negative (60.45/1.0)

TT4 > 53 AND
TT4 <= 136 AND
T3_measured = t AND
referral_source = SVI: sick (168.11/6.11)

sick = f AND

```

```

age <= 70 AND
T4U_measured = t AND
referral_source = other: negative (71.0/15.0)

age <= 32: negative (22.4)

sick = f AND
T4U > 1.12: negative (15.91/0.23)

sick = f AND
referral_source = SVHC: sick (10.04/2.04)

sick = f AND
T4U_measured = f: negative (11.27/1.77)

referral_source = SVI AND
T3_measured = f AND
TT4 > 65: negative (4.94)

sick = f AND
referral_source = SVI AND
T3_measured = t: negative (63.05/20.0)

TT4 <= 137 AND
sick = f AND
referral_source = other AND
TT4 > 62: sick (13.0/1.0)

TT4 > 131: negative (10.48/1.0)

sick = t: sick (5.0)

referral_source = other: negative (4.0/1.0)

: sick (4.1/0.1)

Number of Rules : 17

Time taken to build model: 0.06 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

==== Summary ===

Correctly Classified Instances      1481          98.1445 %
Incorrectly Classified Instances   28           1.8555 %
Kappa statistic                   0.8132
Mean absolute error               0.0209
Root mean squared error          0.1135
Relative absolute error          18.2638 %
Root relative squared error     48.1731 %
Total Number of Instances        1509

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall    F-Measure   MCC      ROC Area   PRC Area   Class
      0.997     0.270     0.983     0.997     0.990     0.820     0.972     0.997   negative
      0.730     0.003     0.942     0.730     0.823     0.820     0.974     0.880     sick
Weighted Avg.      0.981     0.254     0.981     0.981     0.980     0.820     0.972     0.990

==== Confusion Matrix ===

      a     b  <- classified as
1416   4 |     a = negative
      24   65 |     b = sick

```

## Percentage Split 66%

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Remove-R2-5,7-16,18,25-28
Instances:   3772
Attributes:  11
             age
             sick
             TSH_measured
             T3_measured
             T3
             TT4_measured
             TT4

```

```

T4U_measured
T4U
referral_source
Class
Test mode: split 66.0% train, remainder test

==== Classifier model (full training set) ===

PART decision list
-----

T3 > 1.1 AND
TT4 <= 151 AND
TT4 > 56: negative (2986.67)

referral_source = other AND
age <= 72 AND
T4U > 1.02: negative (203.82)

TT4 > 145 AND
T3 > 2.6: negative (117.75)

T3_measured = f AND
referral_source = other: negative (60.45/1.0)

TT4 > 53 AND
TT4 <= 136 AND
T3_measured = t AND
referral_source = SVI: sick (168.11/6.11)

sick = f AND
age <= 70 AND
T4U_measured = t AND
referral_source = other: negative (71.0/15.0)

age <= 32: negative (22.4)

sick = f AND
T4U > 1.12: negative (15.91/0.23)

sick = f AND
referral_source = SVHC: sick (10.04/2.04)

sick = f AND
T4U_measured = f: negative (11.27/1.77)

referral_source = SVI AND
T3_measured = f AND
TT4 > 65: negative (4.94)

sick = f AND
referral_source = SVI AND
T3_measured = t: negative (63.05/20.0)

TT4 <= 137 AND
sick = f AND
referral_source = other AND
TT4 > 62: sick (13.0/1.0)

TT4 > 131: negative (10.48/1.0)

sick = t: sick (5.0)

referral_source = other: negative (4.0/1.0)

: sick (4.1/0.1)

Number of Rules : 17

Time taken to build model: 0.06 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

==== Summary ===

Correctly Classified Instances      1253      97.7379 %
Incorrectly Classified Instances   29       2.2621 %
Kappa statistic                   0.7765
Mean absolute error               0.032
Root mean squared error           0.1412
Relative absolute error           28.0151 %
Root relative squared error      60.1342 %
Total Number of Instances        1282

```

```

==== Detailed Accuracy By Class ====
      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
      0.993    0.280    0.983     0.993    0.988     0.780    0.970     0.996    negative
      0.720    0.007    0.871     0.720    0.788     0.780    0.970     0.760    sick
Weighted Avg.  0.977    0.264    0.976     0.977    0.976     0.780    0.970     0.983

==== Confusion Matrix ====
      a      b  <-- classified as
1199    8 |    a = negative
  21   54 |    b = sick

```

## Leave one out fold

```

==== Run information ====
Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Remove-R2-5,7-16,18,25-28
Instances:   3772
Attributes:  11
              age
              sick
              TSH_measured
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              referral_source
              Class
Test mode:   3772-fold cross-validation

==== Classifier model (full training set) ====
PART decision list
-----
T3 > 1.1 AND
TT4 <= 151 AND
TT4 > 56: negative (2986.67)

referral_source = other AND
age <= 72 AND
T4U > 1.02: negative (203.82)

TT4 > 145 AND
T3 > 2.6: negative (117.75)

T3_measured = f AND
referral_source = other: negative (60.45/1.0)

TT4 > 53 AND
TT4 <= 136 AND
T3_measured = t AND
referral_source = SVI: sick (168.11/6.11)

sick = f AND
age <= 70 AND
T4U_measured = t AND
referral_source = other: negative (71.0/15.0)

age <= 32: negative (22.4)

sick = f AND
T4U > 1.12: negative (15.91/0.23)

sick = f AND
referral_source = SVHC: sick (10.04/2.04)

sick = f AND
T4U_measured = f: negative (11.27/1.77)

referral_source = SVI AND
T3_measured = f AND
TT4 > 65: negative (4.94)

sick = f AND
referral_source = SVI AND
T3_measured = t: negative (63.05/20.0)

TT4 <= 137 AND
sick = f AND

```

```

referral_source = other AND
TT4 > 62: sick (13.0/1.0)

TT4 > 131: negative (10.48/1.0)

sick = t: sick (5.0)

referral_source = other: negative (4.0/1.0)

: sick (4.1/0.1)

Number of Rules : 17

Time taken to build model: 0.07 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3710          98.3563 %
Incorrectly Classified Instances   62           1.6437 %
Kappa statistic                   0.8465
Mean absolute error               0.0247
Root mean squared error          0.1188
Relative absolute error          21.4186 %
Root relative squared error     49.5477 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC    ROC Area   PRC Area   Class
0.996     0.208     0.987     0.996     0.991     0.850     0.982     0.998   negative
0.792     0.004     0.929     0.792     0.855     0.850     0.982     0.862   sick
Weighted Avg.       0.984     0.195     0.983     0.984     0.983     0.850     0.982     0.990

==== Confusion Matrix ====

      a     b  <-- classified as
3527  14 |  a = negative
  48 183 |  b = sick

```

### 3.4.2 Using Wrapper

From the select attribute tab we are choosing **ClassifierSubsetEval** as our classifier with **PART** as the classifier option and the rest to default. Search method we are using **BestFirst** with **bidirectional** as an option.

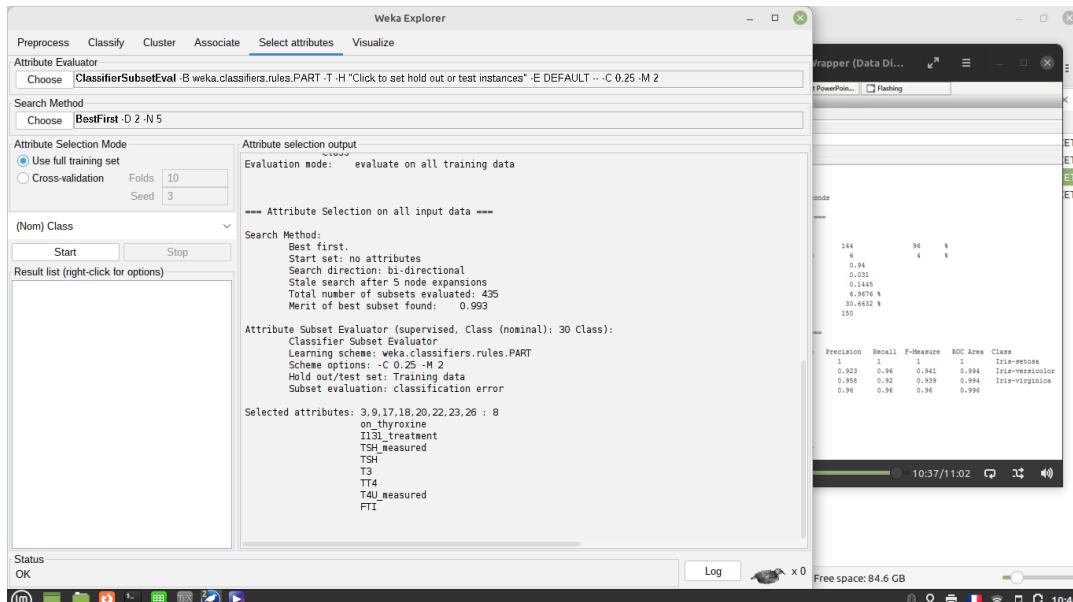


FIGURE 3.4: PART Wrapper1

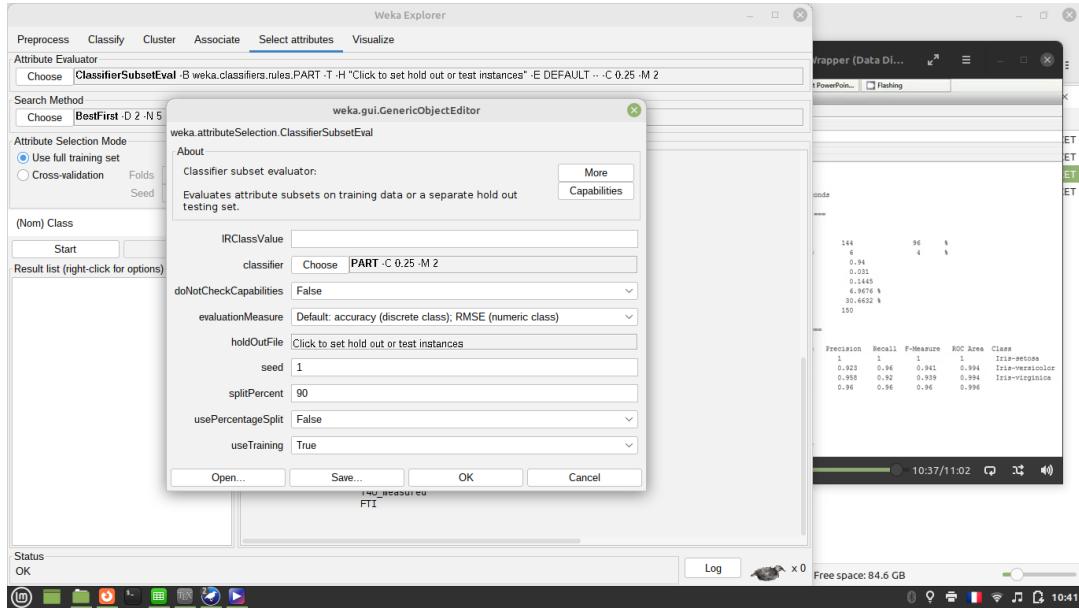


FIGURE 3.5: PART Wrapper2

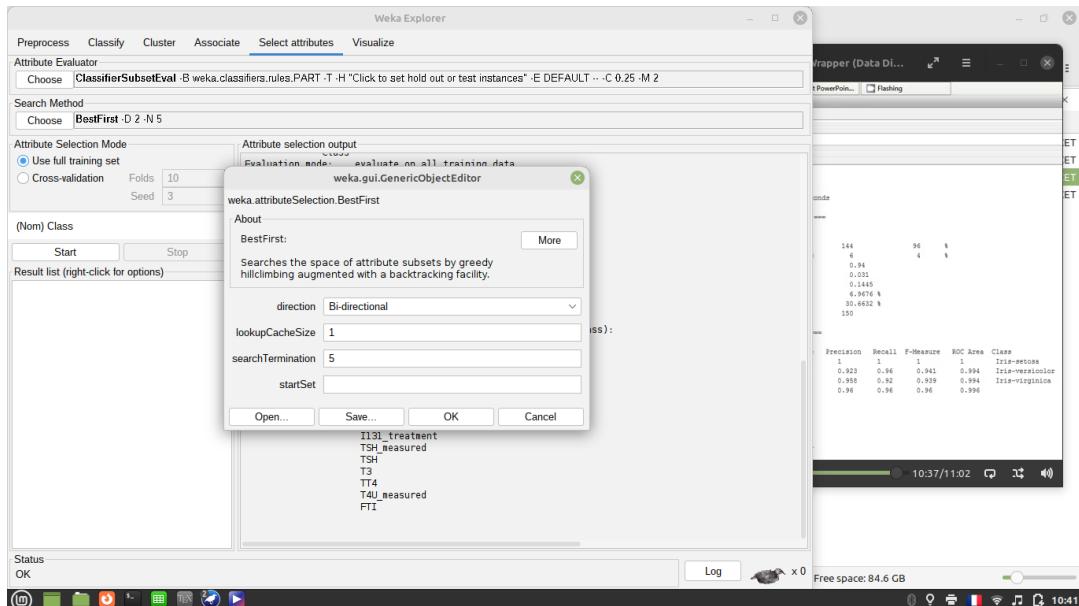


FIGURE 3.6: PART Wrapper3

## Features selection

```
==== Run information ====
Evaluator: weka.attributeSelection.ClassifierSubsetEval -B weka.classifiers.rules.PART -T -H "Click to set hold out or test instances" -E DEFAULT .. -C 0.25 -M 2
Search: weka.attributeSelection.BestFirst -D 2 -N 5
Relation: sick
Instances: 3772
Attributes: 30
age
sex
on_thyroxine
query_on_thyroxine
on_antithyroid_medication
sick
pregnant
thyroid_surgery
I131_treatment
```

```

query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH
T3_measured
T3
TT4_measured
TT4
T4U_measured
T4U
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Evaluation mode: evaluate on all training data

==== Attribute Selection on all input data ====
Search Method:
Best first.
Start set: no attributes
Search direction: bi-directional
Stale search after 5 node expansions
Total number of subsets evaluated: 435
Merit of best subset found: 0.993

Attribute Subset Evaluator (supervised, Class (nominal): 30 Class):
Classifier Subset Evaluator
Learning scheme: weka.classifiers.rules.PART
Scheme options: -C 0.25 -M 2
Hold out/test set: Training data
Subset evaluation: classification error

Selected attributes: 3,9,17,18,20,22,23,26 : 8
on_thyroxine
I131_treatment
TSH_measured
TSH
T3
TT4
T4U_measured
FTI

```

## Cross Validation 5 folds

```

==== Run information ====
Scheme: weka.classifiers.rules.PART -C 0.25 -M 2
Relation: sick-weka.filters.unsupervised.attribute.Remove-R1-2,4-8,10-16,19,21,24-25,27-29
Instances: 3772
Attributes: 9
on_thyroxine
I131_treatment
TSH_measured
TSH
T3
TT4
T4U_measured
FTI
Class
Test mode: 5-fold cross-validation

==== Classifier model (full training set) ====
PART decision list
-----
T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3187.63/1.0)

T3 > 1.1 AND
on_thyroxine = f AND
T3 > 2.6: negative (72.83/0.46)

```

```

TSH_measured = t AND
TSH > 18 AND
TT4 <= 67: negative (75.65)

TSH_measured = f: negative (46.01/2.0)

on_thyroxine = t AND
TSH <= 0.72: negative (43.16)

T3 <= 1.1 AND
TT4 > 140: negative (15.82/1.0)

T4U_measured = t AND
T3 <= 1.1 AND
FTI > 155 AND
TT4 > 92: negative (11.07)

T4U_measured = t AND
T3 <= 1.1 AND
TT4 > 50: sick (234.53/43.81)

on_thyroxine = f AND
T4U_measured = f AND
T3 > 1: negative (9.5)

TSH <= 0.25 AND
TT4 > 151: sick (17.06/1.06)

FTI > 62 AND
T4U_measured = t: negative (28.58/1.0)

on_thyroxine = f AND
T3 <= 1.7 AND
T4U_measured = t AND
TSH <= 3.9: sick (8.45)

TT4 <= 88 AND
T3 > 0.4 AND
TT4 > 42: negative (9.74/0.17)

: sick (11.98/1.77)

Number of Rules : 14

Time taken to build model: 0.09 seconds

==== Stratified cross-validation ====
==== Summary ====

    Correctly Classified Instances      3702           98.1442 %
    Incorrectly Classified Instances     70            1.8558 %
    Kappa statistic                      0.8379
    Mean absolute error                  0.0368
    Root mean squared error              0.1253
    Relative absolute error             31.917 %
    Root relative squared error         52.2457 %
    Total Number of Instances          3772

==== Detailed Accuracy By Class ====

    TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
    0.990     0.156     0.990     0.990     0.990     0.838     0.973     0.997     negative
    0.844     0.010     0.852     0.844     0.848     0.838     0.973     0.809     sick
    Weighted Avg.   0.981     0.147     0.981     0.981     0.981     0.838     0.973     0.985

==== Confusion Matrix ====

    a     b  <- classified as
3507   34 |   a = negative
    36  195 |   b = sick

```

## Cross Validation 10 folds

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Remove-R1-2,4-8,10-16,19,21,24-25,27-29
Instances:   3772
Attributes:  9
              on_thyroxine
              T131_treatment
              TSH_measured
              TSH

```

```

T3
TT4
T4U_measured
FTI
Class
Test mode: 10-fold cross-validation

==== Classifier model (full training set) ====

PART decision list
-----

T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3187.63/1.0)

T3 > 1.1 AND
on_thyroxine = f AND
T3 > 2.6: negative (72.83/0.46)

TSH_measured = t AND
TSH > 18 AND
TT4 <= 67: negative (75.65)

TSH_measured = f: negative (46.01/2.0)

on_thyroxine = t AND
TSH <= 0.72: negative (43.16)

T3 <= 1.1 AND
TT4 > 140: negative (15.82/1.0)

T4U_measured = t AND
T3 <= 1.1 AND
FTI > 155 AND
TT4 > 92: negative (11.07)

T4U_measured = t AND
T3 <= 1.1 AND
TT4 > 50: sick (234.53/43.81)

on_thyroxine = f AND
T4U_measured = f AND
T3 > 1: negative (9.5)

TSH <= 0.25 AND
TT4 > 151: sick (17.06/1.06)

FTI > 62 AND
T4U_measured = t: negative (28.58/1.0)

on_thyroxine = f AND
T3 <= 1.7 AND
T4U_measured = t AND
TSH <= 3.9: sick (8.45)

TT4 <= 88 AND
T3 > 0.4 AND
TT4 > 42: negative (9.74/0.17)

: sick (11.98/1.77)

Number of Rules : 14

Time taken to build model: 0.07 seconds

==== Stratified cross-validation ====
==== Summary ===

  Correctly Classified Instances      3701      98.1177 %
  Incorrectly Classified Instances     71       1.8823 %
  Kappa statistic                      0.8386
  Mean absolute error                  0.036
  Root mean squared error              0.123
  Relative absolute error              31.2687 %
  Root relative squared error          51.3078 %
  Total Number of Instances            3772

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.989     0.139     0.991     0.989     0.990     0.839     0.979     0.997     negative
      0.861     0.011     0.836     0.861     0.849     0.839     0.979     0.860     sick
  Weighted Avg.      0.981     0.131     0.981     0.981     0.981     0.839     0.979     0.989

==== Confusion Matrix ====

```

```

a      b    <-- classified as
3502   39 |     a = negative
 32  199 |     b = sick

```

## Cross Validation 15 folds

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Remove-R1-2,4-8,10-16,19,21,24-25,27-29
Instances:   3772
Attributes:  9
              on_thyroxine
              T131_treatment
              TSH_measured
              TSH
              T3
              TT4
              T4U_measured
              FTI
              Class
Test mode:   15-fold cross-validation

==== Classifier model (full training set) ====

PART decision list
-----
T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3187.63/1.0)

T3 > 1.1 AND
on_thyroxine = f AND
T3 > 2.6: negative (72.83/0.46)

TSH_measured = t AND
TSH > 18 AND
TT4 <= 67: negative (75.65)

TSH_measured = f: negative (46.01/2.0)

on_thyroxine = t AND
TSH <= 0.72: negative (43.16)

T3 <= 1.1 AND
TT4 > 140: negative (15.82/1.0)

T4U_measured = t AND
T3 <= 1.1 AND
FTI > 155 AND
TT4 > 92: negative (11.07)

T4U_measured = t AND
T3 <= 1.1 AND
TT4 > 50: sick (234.53/43.81)

on_thyroxine = f AND
T4U_measured = f AND
T3 > 1: negative (9.5)

TSH <= 0.25 AND
TT4 > 151: sick (17.06/1.06)

FTI > 62 AND
T4U_measured = t: negative (28.58/1.0)

on_thyroxine = f AND
T3 <= 1.7 AND
T4U_measured = t AND
TSH <= 3.9: sick (8.45)

TT4 <= 88 AND
T3 > 0.4 AND
TT4 > 42: negative (9.74/0.17)

: sick (11.98/1.77)

Number of Rules : 14

Time taken to build model: 0.09 seconds

```

```

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3703          98.1707 %
Incorrectly Classified Instances    69           1.8293 %
Kappa statistic                   0.8425
Mean absolute error               0.0353
Root mean squared error           0.1226
Relative absolute error           30.6024 %
Root relative squared error      51.1202 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC     ROC Area   PRC Area   Class
      0.990     0.139     0.991     0.990     0.990     0.843     0.968     0.994     negative
      0.861     0.010     0.843     0.861     0.852     0.843     0.968     0.829     sick
Weighted Avg.       0.982     0.131     0.982     0.982     0.982     0.843     0.968     0.984

==== Confusion Matrix ====

      a      b  <-- classified as
3504   37 |   a = negative
      32  199 |   b = sick

```

## Percentage Split 50%

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Remove-R1-2,4-8,10-16,19,21,24-25,27-29
Instances:   3772
Attributes:  9
              on_thyroxine
              T131_treatment
              TSH_measured
              TSH
              T3
              TT4
              T4U_measured
              FTI
              Class
Test mode:   split 50.0% train, remainder test

==== Classifier model (full training set) ===

PART decision list
-----
T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3187.63/1.0)

T3 > 1.1 AND
on_thyroxine = f AND
T3 > 2.6: negative (72.83/0.46)

TSH_measured = t AND
TSH > 18 AND
TT4 <= 67: negative (75.65)

TSH_measured = f: negative (46.01/2.0)

on_thyroxine = t AND
TSH <= 0.72: negative (43.16)

T3 <= 1.1 AND
TT4 > 140: negative (15.82/1.0)

T4U_measured = t AND
T3 <= 1.1 AND
FTI > 155 AND
TT4 > 92: negative (11.07)

T4U_measured = t AND
T3 <= 1.1 AND
TT4 > 50: sick (234.53/43.81)

on_thyroxine = f AND
T4U_measured = f AND
T3 > 1: negative (9.5)

TSH <= 0.25 AND
TT4 > 151: sick (17.06/1.06)

```

```

FTI > 62 AND
T4U_measured = t: negative (28.58/1.0)

on_thyroxine = f AND
T3 <= 1.7 AND
T4U_measured = t AND
TSH <= 3.9: sick (8.45)

TT4 <= 88 AND
T3 > 0.4 AND
TT4 > 42: negative (9.74/0.17)

: sick (11.98/1.77)

Number of Rules : 14

Time taken to build model: 0.08 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0.01 seconds

==== Summary ===

Correctly Classified Instances      1838          97.4549 %
Incorrectly Classified Instances   48           2.5451 %
Kappa statistic                   0.7684
Mean absolute error               0.0462
Root mean squared error          0.1467
Relative absolute error          40.0654 %
Root relative squared error     60.825 %
Total Number of Instances        1886

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
0.990      0.265      0.983      0.990      0.986      0.770      0.984      0.998   negative
0.735      0.010      0.835      0.735      0.782      0.770      0.984      0.758   sick
Weighted Avg.   0.975      0.249      0.973      0.975      0.974      0.770      0.984      0.983

==== Confusion Matrix ===

      a      b  <- classified as
1752    17 |    a = negative
      31    86 |    b = sick

```

## Percentage Split 60%

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Remove-R1-2,4-8,10-16,19,21,24-25,27-29
Instances:   3772
Attributes:  9
             on_thyroxine
             I131_treatment
             TSH_measured
             TSH
             T3
             TT4
             T4U_measured
             FTI
             Class
Test mode:   split 60.0% train, remainder test

==== Classifier model (full training set) ===

PART decision list
-----
T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3187.63/1.0)

T3 > 1.1 AND
on_thyroxine = f AND
T3 > 2.6: negative (72.83/0.46)

TSH_measured = t AND
TSH > 18 AND
TT4 <= 67: negative (75.65)

```

```

TSH_measured = f: negative (46.01/2.0)

on_thyroxine = t AND
TSH <= 0.72: negative (43.16)

T3 <= 1.1 AND
TT4 > 140: negative (15.82/1.0)

T4U_measured = t AND
T3 <= 1.1 AND
FTI > 155 AND
TT4 > 92: negative (11.07)

T4U_measured = t AND
T3 <= 1.1 AND
TT4 > 50: sick (234.53/43.81)

on_thyroxine = f AND
T4U_measured = f AND
T3 > 1: negative (9.5)

TSH <= 0.25 AND
TT4 > 151: sick (17.06/1.06)

FTI > 62 AND
T4U_measured = t: negative (28.58/1.0)

on_thyroxine = f AND
T3 <= 1.7 AND
T4U_measured = t AND
TSH <= 3.9: sick (8.45)

TT4 <= 88 AND
T3 > 0.4 AND
TT4 > 42: negative (9.74/0.17)

: sick (11.98/1.77)

Number of Rules : 14

Time taken to build model: 0.08 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

==== Summary ===

Correctly Classified Instances      1469          97.3492 %
Incorrectly Classified Instances   40           2.6508 %
Kappa statistic                   0.7331
Mean absolute error               0.0412
Root mean squared error          0.1351
Relative absolute error           35.9378 %
Root relative squared error      57.324  %
Total Number of Instances        1509

==== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC    ROC Area  PRC Area  Class
0.993     0.337    0.979     0.993    0.986     0.740   0.946     0.992   negative
0.663     0.007    0.855     0.663    0.747     0.740   0.946     0.775   sick
Weighted Avg.    0.973    0.318    0.972     0.973    0.972     0.740   0.946     0.980

==== Confusion Matrix ===

      a      b  <- classified as
1410    10 |      a = negative
      30    59 |      b = sick

```

## Percentage Split 66%

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Remove-R1-2,4-8,10-16,19,21,24-25,27-29
Instances:   3772
Attributes:  9
             on_thyroxine
             I131_treatment
             TSH_measured
             TSH
             T3

```

```

TT4
T4U_measured
FTI
Class
Test mode: split 66.0% train, remainder test

==== Classifier model (full training set) ===

PART decision list
-----
T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3187.63/1.0)

T3 > 1.1 AND
on_thyroxine = f AND
T3 > 2.6: negative (72.83/0.46)

TSH_measured = t AND
TSH > 18 AND
TT4 <= 67: negative (75.65)

TSH_measured = f: negative (46.01/2.0)

on_thyroxine = t AND
TSH <= 0.72: negative (43.16)

T3 <= 1.1 AND
TT4 > 140: negative (15.82/1.0)

T4U_measured = t AND
T3 <= 1.1 AND
FTI > 155 AND
TT4 > 92: negative (11.07)

T4U_measured = t AND
T3 <= 1.1 AND
TT4 > 50: sick (234.53/43.81)

on_thyroxine = f AND
T4U_measured = f AND
T3 > 1: negative (9.5)

TSH <= 0.25 AND
TT4 > 151: sick (17.06/1.06)

FTI > 62 AND
T4U_measured = t: negative (28.58/1.0)

on_thyroxine = f AND
T3 <= 1.7 AND
T4U_measured = t AND
TSH <= 3.9: sick (8.45)

TT4 <= 88 AND
T3 > 0.4 AND
TT4 > 42: negative (9.74/0.17)

: sick (11.98/1.77)

Number of Rules : 14

Time taken to build model: 0.08 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0.01 seconds

==== Summary ===

Correctly Classified Instances      1262          98.4399 %
Incorrectly Classified Instances    20           1.5601 %
Kappa statistic                      0.8635
Mean absolute error                  0.0352
Root mean squared error              0.1225
Relative absolute error              30.7984 %
Root relative squared error         52.201  %
Total Number of Instances            1282

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
0.989      0.093     0.994      0.989     0.992      0.864     0.906      0.985    negative
0.907      0.011     0.840      0.907     0.872      0.864     0.906      0.810     sick
Weighted Avg.    0.984     0.089     0.985      0.984     0.985      0.864     0.906      0.974

```

```
==== Confusion Matrix ====
      a     b  <-- classified as
1194   13 |    a = negative
      7   68 |    b = sick
```

## Leave one out fold

```
==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Remove-R1-2,4-8,10-16,19,21,24-25,27-29
Instances:   3772
Attributes:  9
              on_thyroxine
              T3
              TSH_measured
              TSH
              T3
              TT4
              T4U_measured
              FTI
              Class
Test mode:   3772-fold cross-validation

==== Classifier model (full training set) ===

PART decision list
-----
T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3187.63/1.0)

T3 > 1.1 AND
on_thyroxine = f AND
T3 > 2.6: negative (72.83/0.46)

TSH_measured = t AND
TSH > 18 AND
TT4 <= 67: negative (75.65)

TSH_measured = f: negative (46.01/2.0)

on_thyroxine = t AND
TSH <= 0.72: negative (43.16)

T3 <= 1.1 AND
TT4 > 140: negative (15.82/1.0)

T4U_measured = t AND
T3 <= 1.1 AND
FTI > 155 AND
TT4 > 92: negative (11.07)

T4U_measured = t AND
T3 <= 1.1 AND
TT4 > 50: sick (234.53/43.81)

on_thyroxine = f AND
T4U_measured = f AND
T3 > 1: negative (9.5)

TSH <= 0.25 AND
TT4 > 151: sick (17.06/1.06)

FTI > 62 AND
T4U_measured = t: negative (28.58/1.0)

on_thyroxine = f AND
T3 <= 1.7 AND
T4U_measured = t AND
TSH <= 3.9: sick (8.45)

TT4 <= 88 AND
T3 > 0.4 AND
TT4 > 42: negative (9.74/0.17)

: sick (11.98/1.77)

Number of Rules : 14
```

```

Time taken to build model: 0.08 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3676           97.4549 %
Incorrectly Classified Instances   96            2.5451 %
Kappa statistic                   0.7759
Mean absolute error               0.0371
Root mean squared error          0.1356
Relative absolute error          32.2168 %
Root relative squared error     56.5445 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall    F-Measure   MCC      ROC Area   PRC Area   Class
negative  0.987    0.221     0.986     0.987     0.986     0.776    0.957     0.993     negative
          0.779    0.013     0.800     0.779     0.789     0.776    0.957     0.872     sick
Weighted Avg.  0.975    0.208     0.974     0.975     0.974     0.776    0.957     0.986

==== Confusion Matrix ====

      a      b  <-- classified as
3496  45 |  a = negative
      51  180 |  b = sick

```

### 3.4.3 Conclusion

PART							
Preprocessing method	Cross Validation			Percentage Split			Leave One Out Fold
	5 folds	10 folds	15 folds	50.00%	60.00%	66.00%	
With Original Dataset	98.3033%	98.6214%	98.5949%	98.6214%	97.8131%	98.7520%	98.8865%
With Filter	98.1972%	98.2238%	98.0117%	98.0912%	98.1445%	97.7379%	98.3563%
With Wrapper	98.1442%	98.1177%	98.1707%	97.4549%	97.3492%	98.4399%	97.4549%

Feature selection methods didn't help at all with this dataset and PART, out of 12 tests only in one test we saw an improvement and it was with the filter method and Percentage split 60%.

## 3.5 Data Cleaning

### 3.5.1 Missing Values

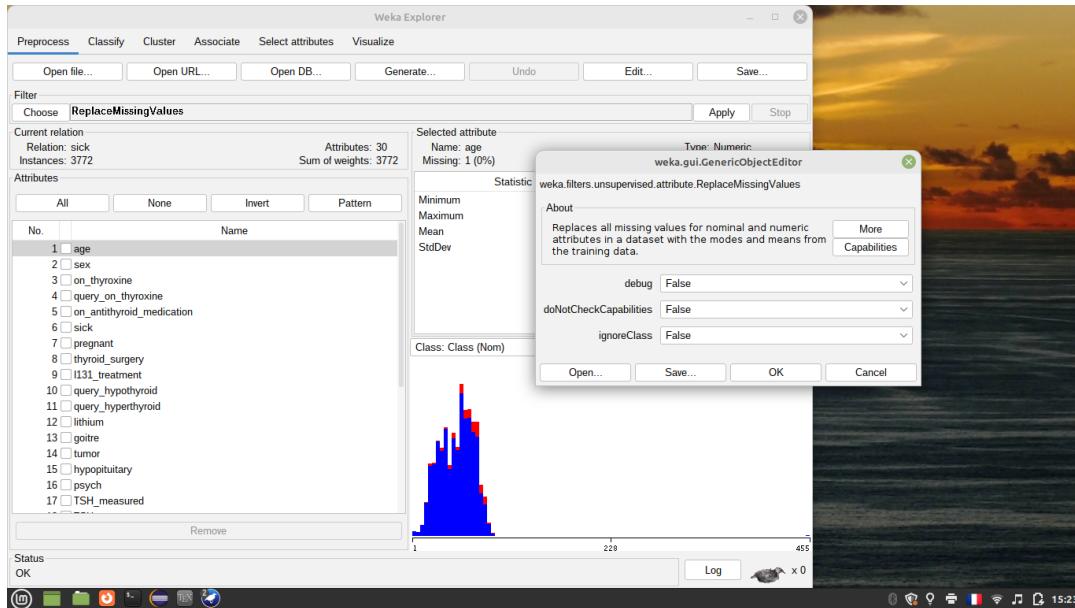


FIGURE 3.7: PART MissingValues

Here we use the **ReplaceMissingValues** filter with our dataset to replace the missing values with the mean value.

### Cross Validation 5 folds

```
== Run information ==

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues
Instances:   3772
Attributes:  30
             age
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             I131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary
             psych
             TSH_measured
             TSH
             T3_measured
             T3
             TT4_measured
             TT4
             T4U_measured
             T4U
             FTI_measured
             FTI
             TBG_measured
             TBG
             referral_source
             Class
Test mode:   5-fold cross-validation

== Classifier model (full training set) ==

PART decision list
-----
T3 > 1.1 AND
FTI <= 171 AND
```

```

TT4 > 56: negative (3281.0/1.0)

referral_source = other AND
sick = f AND
T4U > 1.02: negative (57.0)

T3 > 1.8 AND
referral_source = other AND
on_thyroxine = t: negative (19.0)

T3 > 1.1 AND
T3 > 2.6: negative (49.0)

TT4 <= 55 AND
TSH > 19: negative (47.0)

T3_measured = f AND
sex = F: negative (15.0)

referral_source = SVI AND
FTI <= 157 AND
TT4 > 50 AND
FTI > 64: sick (160.0)

T4U_measured = t AND
sick = f AND
TSH_measured = t AND
T3_measured = t AND
referral_source = other AND
age <= 77 AND
TSH <= 10.3 AND
TT4 <= 156 AND
T3 > 0.4 AND
age <= 71: negative (26.0)

T4U_measured = t AND
TSH_measured = t AND
sick = f AND
T3_measured = t AND
referral_source = other AND
query_hyperthyroid = f AND
query_hypothyroid = f: sick (27.0/5.0)

T4U_measured = t AND
sick = f AND
TSH_measured = t AND
T3_measured = t AND
referral_source = SVHC: sick (9.0/1.0)

T4U_measured = f: negative (7.0)

sick = f AND
TSH_measured = f: negative (6.0/1.0)

sick = f AND
T3_measured = t AND
referral_source = other AND
TT4 <= 172: sick (5.0)

sick = f AND
T3_measured = t AND
referral_source = SVI AND
FTI <= 66 AND
TT4 <= 58: sick (14.0)

sick = f AND
T3_measured = t AND
referral_source = SVI AND
sex = F AND
on_thyroxine = f AND
TSH > 0.17 AND
T4U > 0.57: negative (10.0)

sick = f AND
T3_measured = t AND
referral_source = SVI AND
sex = F AND
on_thyroxine = f: sick (16.0/6.0)

sick = f AND
T3_measured = t: negative (16.0/2.0)

: sick (8.0)

Number of Rules : 18

```

```

Time taken to build model: 0.78 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3701           98.1177 %
Incorrectly Classified Instances    71            1.8823 %
Kappa statistic                   0.8379
Mean absolute error               0.0198
Root mean squared error          0.13
Relative absolute error           17.1641 %
Root relative squared error     54.2213 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.989     0.143     0.991      0.989     0.990      0.838     0.947      0.993     negative
      0.857     0.011     0.839      0.857     0.848      0.838     0.947      0.818     sick
Weighted Avg.      0.981     0.135     0.981      0.981     0.981      0.838     0.947      0.982

==== Confusion Matrix ====

      a      b  <-- classified as
3503  38 |  a = negative
      33  198 |  b = sick

```

## Cross Validation 10 folds

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   10-fold cross-validation

==== Classifier model (full training set) ====

PART decision list
-----
T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3281.0/1.0)

referral_source = other AND
sick = f AND
T4U > 1.02: negative (57.0)

T3 > 1.8 AND
referral_source = other AND
on_thyroxine = t: negative (19.0)

```

```

T3 > 1.1 AND
T3 > 2.6: negative (49.0)

TT4 <= 55 AND
TSH > 19: negative (47.0)

T3_measured = f AND
sex = F: negative (15.0)

referral_source = SVI AND
FTI <= 157 AND
TT4 > 50 AND
FTI > 64: sick (160.0)

T4U_measured = t AND
sick = f AND
TSH_measured = t AND
T3_measured = t AND
referral_source = other AND
age <= 77 AND
TSH <= 10.3 AND
TT4 <= 156 AND
T3 > 0.4 AND
age <= 71: negative (26.0)

T4U_measured = t AND
TSH_measured = t AND
sick = f AND
T3_measured = t AND
referral_source = other AND
query_hyperthyroid = f AND
query_hypothyroid = f: sick (27.0/5.0)

T4U_measured = t AND
sick = f AND
TSH_measured = t AND
T3_measured = t AND
referral_source = SVHC: sick (9.0/1.0)

T4U_measured = f: negative (7.0)

sick = f AND
TSH_measured = f: negative (6.0/1.0)

sick = f AND
T3_measured = t AND
referral_source = other AND
TT4 <= 172: sick (5.0)

sick = f AND
T3_measured = t AND
referral_source = SVI AND
FTI <= 66 AND
TT4 <= 58: sick (14.0)

sick = f AND
T3_measured = t AND
referral_source = SVI AND
sex = F AND
on_thyroxine = f AND
TSH > 0.17 AND
T4U > 0.57: negative (10.0)

sick = f AND
T3_measured = t AND
referral_source = SVI AND
sex = F AND
on_thyroxine = f: sick (16.0/6.0)

sick = f AND
T3_measured = t: negative (16.0/2.0)

: sick (8.0)

Number of Rules : 18

Time taken to build model: 0.31 seconds

==== Stratified cross-validation ====
==== Summary ====

  Correctly Classified Instances      3726          98.7805 %
  Incorrectly Classified Instances     46           1.2195 %
  Kappa statistic                      0.896
  Mean absolute error                  0.0135
  Root mean squared error              0.1056

```

```

Relative absolute error          11.7339 %
Root relative squared error    44.0535 %
Total Number of Instances      3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC     ROC Area   PRC Area   Class
      0.992     0.078     0.995     0.992     0.993     0.896     0.975     0.996   negative
      0.922     0.008     0.884     0.922     0.903     0.896     0.975     0.907   sick
Weighted Avg.      0.988     0.074     0.988     0.988     0.988     0.896     0.975     0.991

==== Confusion Matrix ====

      a      b  <-- classified as
3513  28 |  a = negative
      18 213 |  b = sick

```

## Cross Validation 15 folds

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   15-fold cross-validation

==== Classifier model (full training set) ====

PART decision list
-----
T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3281.0/1.0)

referral_source = other AND
sick = f AND
T4U > 1.02: negative (57.0)

T3 > 1.8 AND
referral_source = other AND
on_thyroxine = t: negative (19.0)

T3 > 1.1 AND
T3 > 2.6: negative (49.0)

TT4 <= 55 AND
TSH > 19: negative (47.0)

T3_measured = f AND
sex = F: negative (15.0)

referral_source = SVI AND

```

```

FTI <= 157 AND
TT4 > 50 AND
FTI > 64: sick (160.0)

T4U_measured = t AND
sick = f AND
TSH_measured = t AND
T3_measured = t AND
referral_source = other AND
age <= 77 AND
TSH <= 10.3 AND
TT4 <= 156 AND
T3 > 0.4 AND
age <= 71: negative (26.0)

T4U_measured = t AND
TSH_measured = t AND
sick = f AND
T3_measured = t AND
referral_source = other AND
query_hyperthyroid = f AND
query_hypothyroid = f: sick (27.0/5.0)

T4U_measured = t AND
sick = f AND
TSH_measured = t AND
T3_measured = t AND
referral_source = SVHC: sick (9.0/1.0)

T4U_measured = f: negative (7.0)

sick = f AND
TSH_measured = f: negative (6.0/1.0)

sick = f AND
T3_measured = t AND
referral_source = other AND
TT4 <= 172: sick (5.0)

sick = f AND
T3_measured = t AND
referral_source = SVI AND
FTI <= 66 AND
FTI <= 58: sick (14.0)

sick = f AND
T3_measured = t AND
referral_source = SVI AND
sex = F AND
on_thyroxine = f AND
TSH > 0.17 AND
T4U > 0.57: negative (10.0)

sick = f AND
T3_measured = t AND
referral_source = SVI AND
sex = F AND
on_thyroxine = f: sick (16.0/6.0)

sick = f AND
T3_measured = t: negative (16.0/2.0)

: sick (8.0)

Number of Rules : 18

Time taken to build model: 0.16 seconds

==== Stratified cross-validation ====
==== Summary ====

  Correctly Classified Instances      3723          98.701 %
  Incorrectly Classified Instances     49           1.299 %
  Kappa statistic                      0.8872
  Mean absolute error                  0.0141
  Root mean squared error              0.1079
  Relative absolute error             12.2782 %
  Root relative squared error        45.0114 %
  Total Number of Instances            3772

==== Detailed Accuracy By Class ====

    TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
    0.993     0.104     0.993     0.993     0.993     0.887     0.959     0.993     negative
    0.896     0.007     0.892     0.896     0.894     0.887     0.959     0.893     sick
  Weighted Avg.     0.987     0.098     0.987     0.987     0.987     0.887     0.959     0.987

```

```
==== Confusion Matrix ====
      a     b  <-- classified as
3516   25 |    a = negative
      24  207 |    b = sick
```

## Percentage Split 50%

```
==== Run information ====
Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 50.0% train, remainder test

==== Classifier model (full training set) ====
PART decision list
-----
T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3281.0/1.0)

referral_source = other AND
sick = f AND
T4U > 1.02: negative (57.0)

T3 > 1.8 AND
referral_source = other AND
on_thyroxine = t: negative (19.0)

T3 > 1.1 AND
T3 > 2.6: negative (49.0)

TT4 <= 55 AND
TSH > 19: negative (47.0)

T3_measured = f AND
sex = F: negative (15.0)

referral_source = SVI AND
FTI <= 157 AND
TT4 > 50 AND
FTI > 64: sick (160.0)

T4U_measured = t AND
sick = f AND
TSH_measured = t AND
T3_measured = t AND
referral_source = other AND
age <= 77 AND
```

```

TSH <= 10.3 AND
TT4 <= 156 AND
T3 > 0.4 AND
age <= 71: negative (26.0)

T4U_measured = t AND
TSH_measured = t AND
sick = f AND
T3_measured = t AND
referral_source = other AND
query_hyperthyroid = f AND
query_hypothyroid = f: sick (27.0/5.0)

T4U_measured = t AND
sick = f AND
TSH_measured = t AND
T3_measured = t AND
referral_source = SVHC: sick (9.0/1.0)

T4U_measured = f: negative (7.0)

sick = f AND
TSH_measured = f: negative (6.0/1.0)

sick = f AND
T3_measured = t AND
referral_source = other AND
TT4 <= 172: sick (5.0)

sick = f AND
T3_measured = t AND
referral_source = SVI AND
FTI <= 66 AND
TT4 <= 58: sick (14.0)

sick = f AND
T3_measured = t AND
referral_source = SVI AND
sex = F AND
on_thyroxine = f AND
TSH > 0.17 AND
T4U > 0.57: negative (10.0)

sick = f AND
T3_measured = t AND
referral_source = SVI AND
sex = F AND
on_thyroxine = f: sick (16.0/6.0)

sick = f AND
T3_measured = t: negative (16.0/2.0)

: sick (8.0)

Number of Rules : 18

Time taken to build model: 0.17 seconds

==== Evaluation on test split ====

Time taken to test model on test split: 0.01 seconds

==== Summary ====
Correctly Classified Instances      1829          96.9777 %
Incorrectly Classified Instances    57           3.0223 %
Kappa statistic                      0.7284
Mean absolute error                  0.0287
Root mean squared error              0.1499
Relative absolute error              24.8669 %
Root relative squared error         62.144  %
Total Number of Instances            1886

==== Detailed Accuracy By Class ====
          TP Rate   FP Rate   Precision   Recall   F-Measure   MCC     ROC Area   PRC Area   Class
          0.987     0.291     0.981     0.987     0.984     0.729     0.942     0.993   negative
          0.709     0.013     0.783     0.709     0.744     0.729     0.942     0.796   sick
Weighted Avg.      0.970     0.273     0.969     0.970     0.969     0.729     0.942     0.980

==== Confusion Matrix ====
      a      b  <- classified as
1746   23 |      a = negative
      34   83 |      b = sick

```

## Percentage Split 60%

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 60.0% train, remainder test

==== Classifier model (full training set) ====

PART decision list
-----
T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3281.0/1.0)

referral_source = other AND
sick = f AND
T4U > 1.02: negative (57.0)

T3 > 1.8 AND
referral_source = other AND
on_thyroxine = t: negative (19.0)

T3 > 1.1 AND
T3 > 2.6: negative (49.0)

TT4 <= 55 AND
TSH > 19: negative (47.0)

T3_measured = f AND
sex = F: negative (15.0)

referral_source = SVI AND
FTI <= 157 AND
TT4 > 50 AND
FTI > 64: sick (160.0)

T4U_measured = t AND
sick = f AND
TSH_measured = t AND
T3_measured = t AND
referral_source = other AND
age <= 77 AND
TSH <= 10.3 AND
TT4 <= 156 AND
T3 > 0.4 AND
age <= 71: negative (26.0)

T4U_measured = t AND
TSH_measured = t AND
sick = f AND
T3_measured = t AND
referral_source = other AND

```

```

query_hyperthyroid = f AND
query_hypothyroid = f: sick (27.0/5.0)

T4U_measured = t AND
sick = f AND
TSH_measured = t AND
T3_measured = t AND
referral_source = SVHC: sick (9.0/1.0)

T4U_measured = f: negative (7.0)

sick = f AND
TSH_measured = f: negative (6.0/1.0)

sick = f AND
T3_measured = t AND
referral_source = other AND
TT4 <= 172: sick (5.0)

sick = f AND
T3_measured = t AND
referral_source = SVI AND
sex = F AND
on_thyroxine = f AND
TSH > 0.17 AND
T4U > 0.57: negative (10.0)

sick = f AND
T3_measured = t AND
referral_source = SVI AND
sex = F AND
on_thyroxine = f: sick (16.0/6.0)

sick = f AND
T3_measured = t: negative (16.0/2.0)

: sick (8.0)

Number of Rules : 18

Time taken to build model: 0.18 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

==== Summary ===

  Correctly Classified Instances      1476          97.8131 %
  Incorrectly Classified Instances    33           2.1869 %
  Kappa statistic                   0.7954
  Mean absolute error               0.023
  Root mean squared error          0.1441
  Relative absolute error          20.0596 %
  Root relative squared error     61.1661 %
  Total Number of Instances        1509

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.991     0.225     0.986     0.991     0.988     0.796     0.917     0.990     negative
      0.775     0.009     0.841     0.775     0.807     0.796     0.917     0.758     sick
  Weighted Avg.      0.978     0.212     0.977     0.978     0.978     0.796     0.917     0.976

==== Confusion Matrix ===

      a      b  <- classified as
  1407   13 |      a = negative
      20   69 |      b = sick

```

## Percentage Split 66%

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues
Instances:   3772

```

```

Attributes: 30
age
sex
on_thyroxine
query_on_thyroxine
on_antithyroid_medication
sick
pregnant
thyroid_surgery
I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH
T3_measured
T3
TT4_measured
TT4
T4U_measured
T4U
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Test mode: split 66.0% train, remainder test

==== Classifier model (full training set) ====

PART decision list
-----
T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3281.0/1.0)

referral_source = other AND
sick = f AND
T4U > 1.02: negative (57.0)

T3 > 1.8 AND
referral_source = other AND
on_thyroxine = t: negative (19.0)

T3 > 1.1 AND
T3 > 2.6: negative (49.0)

TT4 <= 55 AND
TSH > 19: negative (47.0)

T3_measured = f AND
sex = F: negative (15.0)

referral_source = SVI AND
FTI <= 157 AND
TT4 > 50 AND
FTI > 64: sick (160.0)

T4U_measured = t AND
sick = f AND
TSH_measured = t AND
T3_measured = t AND
referral_source = other AND
age <= 77 AND
TSH <= 10.3 AND
TT4 <= 156 AND
T3 > 0.4 AND
age <= 71: negative (26.0)

T4U_measured = t AND
TSH_measured = t AND
sick = f AND
T3_measured = t AND
referral_source = other AND
query_hyperthyroid = f AND
query_hypothyroid = f: sick (27.0/5.0)

T4U_measured = t AND
sick = f AND
TSH_measured = t AND
T3_measured = t AND

```

```

referral_source = SVHC: sick (9.0/1.0)

T4U_measured = f: negative (7.0)

sick = f AND
TSH_measured = f: negative (6.0/1.0)

sick = f AND
T3_measured = t AND
referral_source = other AND
TT4 <= 172: sick (5.0)

sick = f AND
T3_measured = t AND
referral_source = SVI AND
FTI <= 66 AND
FTI <= 58: sick (14.0)

sick = f AND
T3_measured = t AND
referral_source = SVI AND
sex = F AND
on_thyroxine = f AND
TSH > 0.17 AND
T4U > 0.57: negative (10.0)

sick = f AND
T3_measured = t AND
referral_source = SVI AND
sex = F AND
on_thyroxine = f: sick (16.0/6.0)

sick = f AND
T3_measured = t: negative (16.0/2.0)

: sick (8.0)

Number of Rules : 18

Time taken to build model: 0.16 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0.01 seconds

==== Summary ===

Correctly Classified Instances      1269          98.986 %
Incorrectly Classified Instances   13           1.014 %
Kappa statistic                   0.905
Mean absolute error               0.0124
Root mean squared error          0.0908
Relative absolute error          10.8973 %
Root relative squared error     38.6651 %
Total Number of Instances        1282

==== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC    ROC Area  PRC Area  Class
      0.997    0.120    0.993    0.997    0.995    0.906    0.948    0.994    negative
      0.880    0.003    0.943    0.880    0.910    0.906    0.948    0.930    sick
Weighted Avg.      0.990    0.113    0.990    0.990    0.990    0.906    0.948    0.990

==== Confusion Matrix ===

  a     b  <- classified as
1203   4 |  a = negative
   9   66 |  b = sick

```

## Leave one out fold

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick

```

```

pregnant
thyroid_surgery
I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH
T3_measured
T3
TT4_measured
TT4
T4U_measured
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Test mode: 3772-fold cross-validation

== Classifier model (full training set) ==

PART decision list
-----
T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3281.0/1.0)

referral_source = other AND
sick = f AND
T4U > 1.02: negative (57.0)

T3 > 1.8 AND
referral_source = other AND
on_thyroxine = t: negative (19.0)

T3 > 1.1 AND
T3 > 2.6: negative (49.0)

TT4 <= 55 AND
TSH > 19: negative (47.0)

T3_measured = f AND
sex = F: negative (15.0)

referral_source = SVI AND
FTI <= 157 AND
TT4 > 50 AND
FTI > 64: sick (160.0)

T4U_measured = t AND
sick = f AND
TSH_measured = t AND
T3_measured = t AND
referral_source = other AND
age <= 77 AND
TSH <= 10.3 AND
TT4 <= 156 AND
T3 > 0.4 AND
age <= 71: negative (26.0)

T4U_measured = t AND
TSH_measured = t AND
sick = f AND
T3_measured = t AND
referral_source = other AND
query_hyperthyroid = f AND
query_hypothyroid = f: sick (27.0/5.0)

T4U_measured = t AND
sick = f AND
TSH_measured = t AND
T3_measured = t AND
referral_source = SVHC: sick (9.0/1.0)

T4U_measured = f: negative (7.0)

sick = f AND
TSH_measured = f: negative (6.0/1.0)

```

```

sick = f AND
T3_measured = t AND
referral_source = other AND
TT4 <= 172: sick (5.0)

sick = f AND
T3_measured = t AND
referral_source = SVI AND
FTI <= 66 AND
TT4 <= 58: sick (14.0)

sick = f AND
T3_measured = t AND
referral_source = SVI AND
sex = F AND
on_thyroxine = f AND
TSH > 0.17 AND
T4U > 0.57: negative (10.0)

sick = f AND
T3_measured = t AND
referral_source = SVI AND
sex = F AND
on_thyroxine = f: sick (16.0/6.0)

sick = f AND
T3_measured = t: negative (16.0/2.0)

: sick (8.0)

Number of Rules : 18

Time taken to build model: 0.17 seconds

==== Stratified cross-validation ====
==== Summary ====

    Correctly Classified Instances      3725           98.754 %
    Incorrectly Classified Instances   47            1.246 %
    Kappa statistic                   0.8919
    Mean absolute error              0.0137
    Root mean squared error          0.107
    Relative absolute error          11.8944 %
    Root relative squared error     44.6087 %
    Total Number of Instances        3772

==== Detailed Accuracy By Class ====

          TP Rate   FP Rate   Precision   Recall   F-Measure   MCC     ROC Area   PRC Area   Class
          0.993     0.100     0.994     0.993     0.993     0.892     0.956     0.995   negative
          0.900     0.007     0.897     0.900     0.898     0.892     0.956     0.888   sick
Weighted Avg.   0.988     0.094     0.988     0.988     0.988     0.892     0.956     0.988

==== Confusion Matrix ====

      a     b  <- classified as
3517  24 |   a = negative
      23  208 |   b = sick

```

### 3.5.2 Missing Values + Extreme Values

#### Cross Validation 5 folds

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3627
Attributes:  30
             age
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             i131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre

```

```

tumor
hypopituitary
psych
TSH_measured
TSH
T3_measured
T3
TT4_measured
TT4
T4U_measured
T4U
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Test mode: 5-fold cross-validation

==== Classifier model (full training set) ===

PART decision list
-----

T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3230.0/1.0)

T3 > 1.8 AND
referral_source = other AND
on_thyroxine = t: negative (30.0)

T3 > 1.9 AND
referral_source = other AND
query_hyperthyroid = t: negative (15.0)

referral_source = SVI AND
FTI <= 157 AND
TT4 > 53 AND
FTI > 65: sick (157.0)

T3 > 2.6: negative (30.0)

sick = f AND
T4U_measured = t AND
TSH_measured = t AND
referral_source = other AND
query_on_thyroxine = f AND
age <= 77 AND
query_hyperthyroid = f AND
T3_measured = t AND
TT4 <= 156 AND
T3 > 0.8: negative (19.0)

sick = f AND
T4U_measured = t AND
TSH_measured = f: negative (11.0/1.0)

T4U_measured = t AND
sick = f AND
T3_measured = f AND
sex = F: negative (10.0)

T4U_measured = t AND
sick = f AND
T3_measured = t AND
referral_source = other AND
query_on_thyroxine = f AND
age <= 75 AND
query_hypothyroid = f AND
T3 <= 1.7 AND
T3 > 0.5: negative (14.0/2.0)

T4U_measured = t AND
sick = f AND
T3_measured = t AND
referral_source = other AND
query_on_thyroxine = f AND
query_hypothyroid = f AND
age > 40: sick (15.0)

T4U_measured = t AND
sick = f AND
T3_measured = t AND
referral_source = SVHC: sick (9.0/1.0)

T4U_measured = f: negative (7.0)

```

```

sick = t: sick (5.0)

T3_measured = t AND
referral_source = other AND
query_on_thyroxine = f AND
age > 34 AND
age <= 74: negative (4.0)

T3_measured = t AND
referral_source = other AND
query_on_thyroxine = f: sick (5.0)

T3_measured = t AND
referral_source = SVI AND
query_hypothyroid = f AND
sex = M AND
TT4 <= 70 AND
T4U > 0.67: sick (11.0)

T3_measured = t AND
referral_source = SVI AND
sex = F AND
query_hypothyroid = f AND
age > 47 AND
T3 <= 1.9: negative (26.0/6.0)

T3_measured = t AND
sex = F AND
query_hypothyroid = f AND
referral_source = SVI: sick (11.0/1.0)

T3_measured = t: negative (15.0/2.0)

: sick (3.0)

Number of Rules : 20

Time taken to build model: 0.2 seconds

==== Stratified cross-validation ====
==== Summary ====

    Correctly Classified Instances      3562          98.2079 %
    Incorrectly Classified Instances     65           1.7921 %
    Kappa statistic                      0.8463
    Mean absolute error                  0.0191
    Root mean squared error              0.1265
    Relative absolute error              16.2679 %
    Root relative squared error         52.3149 %
    Total Number of Instances            3627

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.991     0.146     0.990     0.991     0.990     0.846     0.953     0.993     negative
      0.854     0.009     0.858     0.854     0.856     0.846     0.953     0.832     sick
Weighted Avg.       0.982     0.138     0.982     0.982     0.982     0.846     0.953     0.983

==== Confusion Matrix ====

      a     b  <- classified as
3369   32 |  a = negative
      33  193 |  b = sick

```

## Cross Validation 10 folds

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3627
Attributes:  30
             age
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             I131_treatment
             query_hypothyroid

```

```
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH
T3_measured
T3
TT4_measured
TT4
T4U_measured
T4U
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Test mode: 10-fold cross-validation

==== Classifier model (full training set) ====

PART decision list
-----
T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3230.0/1.0)

T3 > 1.8 AND
referral_source = other AND
on_thyroxine = t: negative (30.0)

T3 > 1.9 AND
referral_source = other AND
query_hyperthyroid = t: negative (15.0)

referral_source = SVI AND
FTI <= 157 AND
TT4 > 53 AND
FTI > 65: sick (157.0)

T3 > 2.6: negative (30.0)

sick = f AND
T4U_measured = t AND
TSH_measured = t AND
referral_source = other AND
query_on_thyroxine = f AND
age <= 77 AND
query_hyperthyroid = f AND
T3_measured = t AND
T3 <= 156 AND
T3 > 0.8: negative (19.0)

sick = f AND
T4U_measured = t AND
TSH_measured = f: negative (11.0/1.0)

T4U_measured = t AND
sick = f AND
T3_measured = f AND
sex = F: negative (10.0)

T4U_measured = t AND
sick = f AND
T3_measured = t AND
referral_source = other AND
query_on_thyroxine = f AND
age <= 75 AND
query_hypothyroid = f AND
T3 <= 1.7 AND
T3 > 0.5: negative (14.0/2.0)

T4U_measured = t AND
sick = f AND
T3_measured = t AND
referral_source = other AND
query_on_thyroxine = f AND
query_hypothyroid = f AND
age > 40: sick (15.0)

T4U_measured = t AND
sick = f AND
T3_measured = t AND
```

```

referral_source = SVHC: sick (9.0/1.0)

T4U_measured = f: negative (7.0)

sick = t: sick (5.0)

T3_measured = t AND
referral_source = other AND
query_on_thyroxine = f AND
age > 34 AND
age <= 74: negative (4.0)

T3_measured = t AND
referral_source = other AND
query_on_thyroxine = f: sick (5.0)

T3_measured = t AND
referral_source = SVI AND
query_hypothyroid = f AND
sex = M AND
TT4 <= 70 AND
T4U > 0.67: sick (11.0)

T3_measured = t AND
referral_source = SVI AND
sex = F AND
query_hypothyroid = f AND
age > 47 AND
T3 <= 1.9: negative (26.0/6.0)

T3_measured = t AND
sex = F AND
query_hypothyroid = f AND
referral_source = SVI: sick (11.0/1.0)

T3_measured = t: negative (15.0/2.0)

: sick (3.0)

Number of Rules : 20

Time taken to build model: 0.19 seconds

==== Stratified cross-validation ===
==== Summary ===

Correctly Classified Instances      3567           98.3457 %
Incorrectly Classified Instances   60            1.6543 %
Kappa statistic                   0.8596
Mean absolute error               0.0181
Root mean squared error          0.1216
Relative absolute error          15.4883 %
Root relative squared error     50.2879 %
Total Number of Instances        3627

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.991     0.124     0.992     0.991     0.991     0.860     0.962     0.994     negative
      0.876     0.009     0.861     0.876     0.868     0.860     0.962     0.865     sick
Weighted Avg.      0.983     0.117     0.984     0.983     0.984     0.860     0.962     0.986

==== Confusion Matrix ===

      a      b  <- classified as
3369    32 |      a = negative
      28   198 |      b = sick

```

## Cross Validation 15 folds

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3627
Attributes:  30
             age
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant

```

```

thyroid_surgery
I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH
T3_measured
T3
TT4_measured
TT4
T4U_measured
T4U
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Test mode: 15-fold cross-validation

==== Classifier model (full training set) ===

PART decision list
-----
T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3230.0/1.0)

T3 > 1.8 AND
referral_source = other AND
on_thyroxine = t: negative (30.0)

T3 > 1.9 AND
referral_source = other AND
query_hyperthyroid = t: negative (15.0)

referral_source = SVI AND
FTI <= 157 AND
TT4 > 53 AND
FTI > 65: sick (157.0)

T3 > 2.6: negative (30.0)

sick = f AND
T4U_measured = t AND
TSH_measured = t AND
referral_source = other AND
query_on_thyroxine = f AND
age <= 77 AND
query_hyperthyroid = f AND
T3_measured = t AND
TT4 <= 156 AND
T3 > 0.8: negative (19.0)

sick = f AND
T4U_measured = t AND
TSH_measured = f: negative (11.0/1.0)

T4U_measured = t AND
sick = f AND
T3_measured = f AND
sex = F: negative (10.0)

T4U_measured = t AND
sick = f AND
T3_measured = t AND
referral_source = other AND
query_on_thyroxine = f AND
age <= 75 AND
query_hypothyroid = f AND
T3 <= 1.7 AND
T3 > 0.5: negative (14.0/2.0)

T4U_measured = t AND
sick = f AND
T3_measured = t AND
referral_source = other AND
query_on_thyroxine = f AND
query_hypothyroid = f AND
age > 40: sick (15.0)

```

```

T4U_measured = t AND
sick = f AND
T3_measured = t AND
referral_source = SVHC: sick (9.0/1.0)

T4U_measured = f: negative (7.0)

sick = t: sick (5.0)

T3_measured = t AND
referral_source = other AND
query_on_thyroxine = f AND
age > 34 AND
age <= 74: negative (4.0)

T3_measured = t AND
referral_source = other AND
query_on_thyroxine = f: sick (5.0)

T3_measured = t AND
referral_source = SVI AND
query_hypothyroid = f AND
sex = M AND
TT4 <= 70 AND
T4U > 0.67: sick (11.0)

T3_measured = t AND
referral_source = SVI AND
sex = F AND
query_hypothyroid = f AND
age > 47 AND
T3 <= 1.9: negative (26.0/6.0)

T3_measured = t AND
sex = F AND
query_hypothyroid = f AND
referral_source = SVI: sick (11.0/1.0)

T3_measured = t: negative (15.0/2.0)

: sick (3.0)

Number of Rules : 20

Time taken to build model: 0.13 seconds

==== Stratified cross-validation ====
==== Summary ====

  Correctly Classified Instances      3566           98.3182 %
  Incorrectly Classified Instances    61            1.6818 %
  Kappa statistic                   0.8558
  Mean absolute error               0.0176
  Root mean squared error          0.1208
  Relative absolute error          15.0298 %
  Root relative squared error     49.9718 %
  Total Number of Instances        3627

==== Detailed Accuracy By Class ====

          TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
          0.991     0.137     0.991     0.991     0.991     0.856     0.958     0.994     negative
          0.863     0.009     0.867     0.863     0.865     0.856     0.958     0.877     sick
  Weighted Avg.       0.983     0.129     0.983     0.983     0.983     0.856     0.958     0.987

==== Confusion Matrix ====

      a      b  <-- classified as
3371   30 |   a = negative
      31  195 |   b = sick

```

## Percentage Split 50%

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3627
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine

```

```

on_antithyroid_medication
sick
pregnant
thyroid_surgery
I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH
T3_measured
T3
TT4_measured
TT4
T4U_measured
T4U
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Test mode: split 50.0% train, remainder test

```

```
==== Classifier model (full training set) ====
```

```
PART decision list
```

```
-----
T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3230.0/1.0)
```

```
T3 > 1.8 AND
referral_source = other AND
on_thyroxine = t: negative (30.0)
```

```
T3 > 1.9 AND
referral_source = other AND
query_hyperthyroid = t: negative (15.0)
```

```
referral_source = SVI AND
FTI <= 157 AND
TT4 > 53 AND
FTI > 65: sick (157.0)
```

```
T3 > 2.6: negative (30.0)
```

```
sick = f AND
T4U_measured = t AND
TSH_measured = t AND
referral_source = other AND
query_on_thyroxine = f AND
age <= 77 AND
query_hyperthyroid = f AND
T3_measured = t AND
TT4 <= 156 AND
T3 > 0.8: negative (19.0)
```

```
sick = f AND
T4U_measured = t AND
TSH_measured = f: negative (11.0/1.0)
```

```
T4U_measured = t AND
sick = f AND
T3_measured = f AND
sex = F: negative (10.0)
```

```
T4U_measured = t AND
sick = f AND
T3_measured = t AND
referral_source = other AND
query_on_thyroxine = f AND
age <= 75 AND
query_hypothyroid = f AND
T3 <= 1.7 AND
T3 > 0.5: negative (14.0/2.0)
```

```
T4U_measured = t AND
sick = f AND
T3_measured = t AND
referral_source = other AND
query_on_thyroxine = f AND
```

```

query_hypothyroid = f AND
age > 40: sick (15.0)

T4U_measured = t AND
sick = f AND
T3_measured = t AND
referral_source = SVHC: sick (9.0/1.0)

T4U_measured = f: negative (7.0)

sick = t: sick (5.0)

T3_measured = t AND
referral_source = other AND
query_on_thyroxine = f AND
age > 34 AND
age <= 74: negative (4.0)

T3_measured = t AND
referral_source = other AND
query_on_thyroxine = f: sick (5.0)

T3_measured = t AND
referral_source = SVI AND
query_hypothyroid = f AND
sex = M AND
TT4 <= 70 AND
T4U > 0.67: sick (11.0)

T3_measured = t AND
referral_source = SVI AND
sex = F AND
query_hypothyroid = f AND
age > 47 AND
T3 <= 1.9: negative (26.0/6.0)

T3_measured = t AND
sex = F AND
query_hypothyroid = f AND
referral_source = SVI: sick (11.0/1.0)

T3_measured = t: negative (15.0/2.0)

: sick (3.0)

Number of Rules : 20

Time taken to build model: 0.11 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0.01 seconds

==== Summary ===

    Correctly Classified Instances      1781          98.235 %
    Incorrectly Classified Instances     32           1.765 %
    Kappa statistic                      0.8383
    Mean absolute error                  0.0196
    Root mean squared error              0.1212
    Relative absolute error             16.6718 %
    Root relative squared error         51.6293 %
    Total Number of Instances           1813

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.991     0.160     0.990      0.991     0.991     0.838     0.964     0.996     negative
      0.840     0.009     0.856      0.840     0.848     0.838     0.964     0.864     sick
Weighted Avg.     0.982     0.152     0.982      0.982     0.982     0.838     0.964     0.988

==== Confusion Matrix ===

      a     b  <-- classified as
1692   15 |     a = negative
      17   89 |     b = sick
  
```

## Percentage Split 60%

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
  
```

```

Instances:      3627
Attributes:    30
               age
               sex
               on_thyroxine
               query_on_thyroxine
               on_antithyroid_medication
               sick
               pregnant
               thyroid_surgery
               T131_treatment
               query_hypothyroid
               query_hyperthyroid
               lithium
               goitre
               tumor
               hypopituitary
               psych
               TSH_measured
               TSH
               T3_measured
               T3
               TT4_measured
               TT4
               T4U_measured
               T4U
               FTI_measured
               FTI
               TBG_measured
               TBG
               referral_source
               Class
Test mode:     split 60.0% train, remainder test

==== Classifier model (full training set) ====

PART decision list
-----
T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3230.0/1.0)

T3 > 1.8 AND
referral_source = other AND
on_thyroxine = t: negative (30.0)

T3 > 1.9 AND
referral_source = other AND
query_hyperthyroid = t: negative (15.0)

referral_source = SVI AND
FTI <= 157 AND
TT4 > 53 AND
FTI > 65: sick (157.0)

T3 > 2.6: negative (30.0)

sick = f AND
T4U_measured = t AND
TSH_measured = t AND
referral_source = other AND
query_on_thyroxine = f AND
age <= 77 AND
query_hyperthyroid = f AND
T3_measured = t AND
TT4 <= 156 AND
T3 > 0.8: negative (19.0)

sick = f AND
T4U_measured = t AND
TSH_measured = f: negative (11.0/1.0)

T4U_measured = t AND
sick = f AND
T3_measured = f AND
sex = F: negative (10.0)

T4U_measured = t AND
sick = f AND
T3_measured = t AND
referral_source = other AND
query_on_thyroxine = f AND
age <= 75 AND
query_hypothyroid = f AND
T3 <= 1.7 AND
T3 > 0.5: negative (14.0/2.0)

```

```

T4U_measured = t AND
sick = f AND
T3_measured = t AND
referral_source = other AND
query_on_thyroxine = f AND
query_hypothyroid = f AND
age > 40: sick (15.0)

T4U_measured = t AND
sick = f AND
T3_measured = t AND
referral_source = SVHC: sick (9.0/1.0)

T4U_measured = f: negative (7.0)

sick = t: sick (5.0)

T3_measured = t AND
referral_source = other AND
query_on_thyroxine = f AND
age > 34 AND
age <= 74: negative (4.0)

T3_measured = t AND
referral_source = other AND
query_on_thyroxine = f: sick (5.0)

T3_measured = t AND
referral_source = SVI AND
query_hypothyroid = f AND
sex = M AND
TT4 <= 70 AND
T4U > 0.67: sick (11.0)

T3_measured = t AND
referral_source = SVI AND
sex = F AND
query_hypothyroid = f AND
age > 47 AND
T3 <= 1.9: negative (26.0/6.0)

T3_measured = t AND
sex = F AND
query_hypothyroid = f AND
referral_source = SVI: sick (11.0/1.0)

T3_measured = t: negative (15.0/2.0)

: sick (3.0)

Number of Rules : 20

Time taken to build model: 0.12 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

==== Summary ===

    Correctly Classified Instances      1427           98.346 %
    Incorrectly Classified Instances     24            1.654 %
    Kappa statistic                      0.8484
    Mean absolute error                  0.0171
    Root mean squared error              0.1271
    Relative absolute error             14.8327 %
    Root relative squared error         54.9805 %
    Total Number of Instances          1451

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.990     0.122     0.993     0.990     0.991     0.849     0.946     0.994     negative
      0.878     0.010     0.837     0.878     0.857     0.849     0.946     0.773     sick
Weighted Avg.      0.983     0.116     0.984     0.983     0.984     0.849     0.946     0.981

==== Confusion Matrix ===

      a      b  <-- classified as
1355    14 |      a = negative
      10    72 |      b = sick

```

## Percentage Split 66%

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3627
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 66.0% train, remainder test

==== Classifier model (full training set) ====

PART decision list
-----
T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3230.0/1.0)

T3 > 1.8 AND
referral_source = other AND
on_thyroxine = t: negative (30.0)

T3 > 1.9 AND
referral_source = other AND
query_hyperthyroid = t: negative (15.0)

referral_source = SVI AND
FTI <= 157 AND
TT4 > 53 AND
FTI > 65: sick (157.0)

T3 > 2.6: negative (30.0)

sick = f AND
T4U_measured = t AND
TSH_measured = t AND
referral_source = other AND
query_on_thyroxine = f AND
age <= 77 AND
query_hyperthyroid = f AND
T3_measured = t AND
TT4 <= 156 AND
T3 > 0.8: negative (19.0)

sick = f AND
T4U_measured = t AND
TSH_measured = f: negative (11.0/1.0)

T4U_measured = t AND
sick = f AND
T3_measured = f AND
sex = F: negative (10.0)

T4U_measured = t AND
sick = f AND
T3_measured = t AND

```

```

referral_source = other AND
query_on_thyroxine = f AND
age <= 75 AND
query_hypothyroid = f AND
T3 <= 1.7 AND
T3 > 0.5: negative (14.0/2.0)

T4U_measured = t AND
sick = f AND
T3_measured = t AND
referral_source = other AND
query_on_thyroxine = f AND
query_hypothyroid = f AND
age > 40: sick (15.0)

T4U_measured = t AND
sick = f AND
T3_measured = t AND
referral_source = SVHC: sick (9.0/1.0)

T4U_measured = f: negative (7.0)

sick = t: sick (5.0)

T3_measured = t AND
referral_source = other AND
query_on_thyroxine = f AND
age > 34 AND
age <= 74: negative (4.0)

T3_measured = t AND
referral_source = other AND
query_on_thyroxine = f: sick (5.0)

T3_measured = t AND
referral_source = SVI AND
query_hypothyroid = f AND
sex = M AND
TT4 <= 70 AND
T4U > 0.67: sick (11.0)

T3_measured = t AND
referral_source = SVI AND
sex = F AND
query_hypothyroid = f AND
age > 47 AND
T3 <= 1.9: negative (26.0/6.0)

T3_measured = t AND
sex = F AND
query_hypothyroid = f AND
referral_source = SVI: sick (11.0/1.0)

T3_measured = t: negative (15.0/2.0)

: sick (3.0)

Number of Rules : 20

Time taken to build model: 0.12 seconds

==== Evaluation on test split ====

Time taken to test model on test split: 0 seconds

==== Summary ====

Correctly Classified Instances      1211      98.2157 %
Incorrectly Classified Instances    22       1.7843 %
Kappa statistic                      0.8476
Mean absolute error                  0.0184
Root mean squared error              0.1261
Relative absolute error              15.7549 %
Root relative squared error         52.7676 %
Total Number of Instances           1233

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
          0.989     0.120     0.992     0.989     0.990     0.848     0.969     0.996   negative
          0.880     0.011     0.835     0.880     0.857     0.848     0.969     0.835     sick
Weighted Avg.      0.982     0.113     0.983     0.982     0.982     0.848     0.969     0.986

==== Confusion Matrix ====

  a     b  <-- classified as

```

```
1145 13 | a = negative
9 66 | b = sick
```

## Leave one out fold

```
==== Run information ===

Scheme: weka.classifiers.rules.PART -C 0.25 -M 2
Relation: sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances: 3627
Attributes: 30
age
sex
on_thyroxine
query_on_thyroxine
on_antithyroid_medication
sick
pregnant
thyroid_surgery
I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH
T3_measured
T3
TT4_measured
TT4
T4U_measured
T4U
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Test mode: 3627-fold cross-validation

==== Classifier model (full training set) ===

PART decision list
-----
T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3230.0/1.0)

T3 > 1.8 AND
referral_source = other AND
on_thyroxine = t: negative (30.0)

T3 > 1.9 AND
referral_source = other AND
query_hyperthyroid = t: negative (15.0)

referral_source = SVI AND
FTI <= 157 AND
TT4 > 53 AND
FTI > 65: sick (157.0)

T3 > 2.6: negative (30.0)

sick = f AND
T4U_measured = t AND
TSH_measured = t AND
referral_source = other AND
query_on_thyroxine = f AND
age <= 77 AND
query_hyperthyroid = f AND
T3_measured = t AND
T3 <= 156 AND
T3 > 0.8: negative (19.0)

sick = f AND
T4U_measured = t AND
TSH_measured = f: negative (11.0/1.0)

T4U_measured = t AND
sick = f AND
```

```

T3_measured = f AND
sex = F: negative (10.0)

T4U_measured = t AND
sick = f AND
T3_measured = t AND
referral_source = other AND
query_on_thyroxine = f AND
age <= 75 AND
query_hypothyroid = f AND
T3 <= 1.7 AND
T3 > 0.5: negative (14.0/2.0)

T4U_measured = t AND
sick = f AND
T3_measured = t AND
referral_source = other AND
query_on_thyroxine = f AND
query_hypothyroid = f AND
age > 40: sick (15.0)

T4U_measured = t AND
sick = f AND
T3_measured = t AND
referral_source = SVHC: sick (9.0/1.0)

T4U_measured = f: negative (7.0)

sick = t: sick (5.0)

T3_measured = t AND
referral_source = other AND
query_on_thyroxine = f AND
age > 34 AND
age <= 74: negative (4.0)

T3_measured = t AND
referral_source = other AND
query_on_thyroxine = f: sick (5.0)

T3_measured = t AND
referral_source = SVI AND
query_hypothyroid = f AND
sex = M AND
TT4 <= 70 AND
T4U > 0.67: sick (11.0)

T3_measured = t AND
referral_source = SVI AND
sex = F AND
query_hypothyroid = f AND
age > 47 AND
T3 <= 1.9: negative (26.0/6.0)

T3_measured = t AND
sex = F AND
query_hypothyroid = f AND
referral_source = SVI: sick (11.0/1.0)

T3_measured = t: negative (15.0/2.0)

: sick (3.0)

Number of Rules : 20

Time taken to build model: 0.12 seconds

==== Stratified cross-validation ====
==== Summary ====

  Correctly Classified Instances      3567      98.3457 %
  Incorrectly Classified Instances     60       1.6543 %
  Kappa statistic                      0.8578
  Mean absolute error                  0.0178
  Root mean squared error              0.1208
  Relative absolute error              15.2343 %
  Root relative squared error          49.9636 %
  Total Number of Instances            3627

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.991     0.137     0.991      0.991     0.991     0.858     0.945      0.993    negative
      0.863     0.009     0.871      0.863     0.867     0.858     0.945      0.879    sick
  Weighted Avg.      0.983     0.129     0.983      0.983     0.983     0.858     0.945      0.986

```

```
==== Confusion Matrix ====
      a     b  <-- classified as
3372   29 |    a = negative
      31  195 |    b = sick
```

### 3.5.3 Missing Values + Outliers

#### Cross Validation 5 folds

```
==== Run information ====
Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3548
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   5-fold cross-validation

==== Classifier model (full training set) ====
PART decision list
-----
T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3157.0/1.0)

referral_source = SVI AND
TSH <= 13 AND
FTI <= 157 AND
TT4 > 50 AND
FTI > 64: sick (157.0)

sick = f AND
T4U_measured = t AND
referral_source = other AND
age <= 77 AND
query_hypothyroid = f AND
T3_measured = t AND
TT4 <= 58: negative (31.0)

sick = f AND
T4U_measured = t AND
T3 > 2 AND
on_thyroxine = t: negative (24.0)

sick = f AND
T4U_measured = t AND
T3 <= 2.8 AND
TSH_measured = t AND
referral_source = SVI AND
TSH > 9.8: negative (20.0)
```

```

sick = f AND
T4U_measured = t AND
T3 > 2.8: negative (14.0)

sick = f AND
T4U_measured = f: negative (12.0)

sick = f AND
TSH_measured = f: negative (11.0/1.0)

sick = f AND
referral_source = SVHC: sick (9.0/1.0)

sick = f AND
referral_source = other AND
query_on_thyroxine = f AND
query_hyperthyroid = f AND
T3_measured = t AND
query_hypothyroid = t AND
age > 77: sick (9.0/1.0)

sick = f AND
referral_source = other AND
query_on_thyroxine = f AND
query_hyperthyroid = f AND
T3_measured = t AND
query_hypothyroid = t AND
T3 > 0.9: negative (7.0)

sick = f AND
FTI <= 66: sick (19.0/2.0)

sick = f AND
query_hypothyroid = f AND
T3_measured = t AND
TSH > 0.025 AND
referral_source = SVI AND
sex = F AND
FTI <= 176: negative (11.0/1.0)

sick = f AND
T3_measured = f: negative (8.0)

sick = f AND
query_hypothyroid = f AND
TSH > 0.025 AND
referral_source = SVI AND
sex = M: negative (6.0)

sick = f AND
query_hypothyroid = f AND
referral_source = other AND
TT4 <= 155 AND
T4U > 0.56 AND
TSH <= 34: negative (21.0/1.0)

TT4 <= 161: sick (21.0)

TSH <= 0.32 AND
referral_source = other AND
age > 54: negative (3.0/1.0)

TSH <= 0.32: sick (5.0)

: negative (3.0)

Number of Rules : 20

Time taken to build model: 0.15 seconds

==== Stratified cross-validation ====
==== Summary ====

  Correctly Classified Instances      3488      98.3089 %
  Incorrectly Classified Instances     60       1.6911 %
  Kappa statistic                      0.8576
  Mean absolute error                  0.0179
  Root mean squared error              0.1239
  Relative absolute error              15.2603 %
  Root relative squared error          51.2789 %
  Total Number of Instances            3548

==== Detailed Accuracy By Class ====

           TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
           0.990     0.118     0.992     0.990     0.991     0.858     0.937     0.991   negative
           0.882     0.010     0.852     0.882     0.867     0.858     0.937     0.848    sick
  Weighted Avg.      0.983     0.111     0.983     0.983     0.983     0.858     0.937     0.983

```

```
==== Confusion Matrix ====
      a     b  <-- classified as
3293   34 |    a = negative
      26  195 |    b = sick
```

## Cross Validation 10 folds

```
==== Run information ====
Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3548
Attributes:  30
            age
            sex
            on_thyroxine
            query_on_thyroxine
            on_antithyroid_medication
            sick
            pregnant
            thyroid_surgery
            t131_treatment
            query_hypothyroid
            query_hyperthyroid
            lithium
            goitre
            tumor
            hypopituitary
            psych
            TSH_measured
            TSH
            T3_measured
            T3
            TT4_measured
            TT4
            T4U_measured
            T4U
            FTI_measured
            FTI
            TBG_measured
            TBG
            referral_source
            Class
Test mode:   10-fold cross-validation
==== Classifier model (full training set) ====
PART decision list
-----
T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3157.0/1.0)

referral_source = SVI AND
TSH <= 13 AND
FTI <= 157 AND
TT4 > 50 AND
FTI > 64: sick (157.0)

sick = f AND
T4U_measured = t AND
referral_source = other AND
age <= 77 AND
query_hypothyroid = f AND
T3_measured = t AND
TT4 <= 58: negative (31.0)

sick = f AND
T4U_measured = t AND
T3 > 2 AND
on_thyroxine = t: negative (24.0)

sick = f AND
T4U_measured = t AND
T3 <= 2.8 AND
TSH_measured = t AND
referral_source = SVI AND
TSH > 9.8: negative (20.0)

sick = f AND
T4U_measured = t AND
T3 > 2.8: negative (14.0)
```

```

sick = f AND
T4U_measured = f: negative (12.0)

sick = f AND
TSH_measured = f: negative (11.0/1.0)

sick = f AND
referral_source = SVHC: sick (9.0/1.0)

sick = f AND
referral_source = other AND
query_on_thyroxine = f AND
age > 77: sick (9.0/1.0)

sick = f AND
referral_source = other AND
query_on_thyroxine = f AND
query_hyperthyroid = f AND
T3_measured = t AND
query_hypothyroid = t AND
T3 > 0.9: negative (7.0)

sick = f AND
FTI <= 66: sick (19.0/2.0)

sick = f AND
query_hypothyroid = f AND
T3_measured = t AND
TSH > 0.025 AND
referral_source = SVI AND
sex = F AND
FTI <= 176: negative (11.0/1.0)

sick = f AND
T3_measured = f: negative (8.0)

sick = f AND
query_hypothyroid = f AND
TSH > 0.025 AND
referral_source = SVI AND
sex = M: negative (6.0)

sick = f AND
query_hypothyroid = f AND
referral_source = other AND
TT4 <= 155 AND
T4U > 0.56 AND
TSH <= 34: negative (21.0/1.0)

TT4 <= 161: sick (21.0)

TSH <= 0.32 AND
referral_source = other AND
age > 54: negative (3.0/1.0)

TSH <= 0.32: sick (5.0)

: negative (3.0)

Number of Rules : 20

Time taken to build model: 0.13 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3497          98.5626 %
Incorrectly Classified Instances    51           1.4374 %
Kappa statistic                   0.8777
Mean absolute error               0.0158
Root mean squared error          0.1157
Relative absolute error           13.5115 %
Root relative squared error      47.8541 %
Total Number of Instances         3548

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
          0.992     0.109     0.993     0.992     0.992     0.878     0.955     0.993   negative
          0.891     0.008     0.879     0.891     0.885     0.878     0.955     0.864     sick
Weighted Avg.      0.986     0.102     0.986     0.986     0.986     0.878     0.955     0.985

==== Confusion Matrix ====

  a      b  <-- classified as

```

```
3300 27 | a = negative
24 197 | b = sick
```

## Cross Validation 15 folds

```
==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3548
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   15-fold cross-validation

==== Classifier model (full training set) ===

PART decision list
-----

T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3157.0/1.0)

referral_source = SVI AND
TSH <= 13 AND
FTI <= 157 AND
TT4 > 50 AND
FTI > 64: sick (157.0)

sick = f AND
T4U_measured = t AND
referral_source = other AND
age <= 77 AND
query_hypothyroid = f AND
T3_measured = t AND
TT4 <= 58: negative (31.0)

sick = f AND
T4U_measured = t AND
T3 > 2 AND
on_thyroxine = t: negative (24.0)

sick = f AND
T4U_measured = t AND
T3 <= 2.8 AND
TSH_measured = t AND
referral_source = SVI AND
TSH > 9.8: negative (20.0)

sick = f AND
T4U_measured = t AND
T3 > 2.8: negative (14.0)

sick = f AND
T4U_measured = f: negative (12.0)
```

```

sick = f AND
TSH_measured = f: negative (11.0/1.0)

sick = f AND
referral_source = SVHC: sick (9.0/1.0)

sick = f AND
referral_source = other AND
query_on_thyroxine = f AND
age > 77: sick (9.0/1.0)

sick = f AND
referral_source = other AND
query_on_thyroxine = f AND
query_hyperthyroid = f AND
T3_measured = t AND
query_hypothyroid = t AND
T3 > 0.9: negative (7.0)

sick = f AND
FTI <= 66: sick (19.0/2.0)

sick = f AND
query_hypothyroid = f AND
T3_measured = t AND
TSH > 0.025 AND
referral_source = SVI AND
sex = F AND
FTI <= 176: negative (11.0/1.0)

sick = f AND
T3_measured = f: negative (8.0)

sick = f AND
query_hypothyroid = f AND
TSH > 0.025 AND
referral_source = SVI AND
sex = M: negative (6.0)

sick = f AND
query_hypothyroid = f AND
referral_source = other AND
TT4 <= 155 AND
T4U > 0.56 AND
TSH <= 34: negative (21.0/1.0)

TT4 <= 161: sick (21.0)

TSH <= 0.32 AND
referral_source = other AND
age > 54: negative (3.0/1.0)

TSH <= 0.32: sick (5.0)

: negative (3.0)

Number of Rules : 20

Time taken to build model: 0.12 seconds

==== Stratified cross-validation ====
==== Summary ====

  Correctly Classified Instances      3495          98.5062 %
  Incorrectly Classified Instances     53           1.4938 %
  Kappa statistic                      0.8713
  Mean absolute error                  0.0161
  Root mean squared error              0.1183
  Relative absolute error             13.7447 %
  Root relative squared error        48.9502 %
  Total Number of Instances            3548

==== Detailed Accuracy By Class ====

          TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
  0.992      0.127      0.992      0.992      0.992      0.871      0.942      0.992   negative
  0.873      0.008      0.885      0.873      0.879      0.871      0.942      0.862    sick
  Weighted Avg.      0.985      0.119      0.985      0.985      0.985      0.871      0.942      0.984

==== Confusion Matrix ====

      a      b  <- classified as
3302   25 |      a = negative
      28   193 |      b = sick

```

## Percentage Split 50%

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3548
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 50.0% train, remainder test

==== Classifier model (full training set) ===

PART decision list
-----
T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3157.0/1.0)

referral_source = SVI AND
TSH <= 13 AND
FTI <= 157 AND
TT4 > 50 AND
FTI > 64: sick (157.0)

sick = f AND
T4U_measured = t AND
referral_source = other AND
age <= 77 AND
query_hypothyroid = f AND
T3_measured = t AND
TT4 <= 58: negative (31.0)

sick = f AND
T4U_measured = t AND
T3 > 2 AND
on_thyroxine = t: negative (24.0)

sick = f AND
T4U_measured = t AND
T3 <= 2.8 AND
TSH_measured = t AND
referral_source = SVI AND
TSH > 9.8: negative (20.0)

sick = f AND
T4U_measured = t AND
T3 > 2.8: negative (14.0)

sick = f AND
T4U_measured = f: negative (12.0)

sick = f AND
TSH_measured = f: negative (11.0/1.0)

sick = f AND
referral_source = SVHC: sick (9.0/1.0)

```

```

sick = f AND
referral_source = other AND
query_on_thyroxine = f AND
age > 77: sick (9.0/1.0)

sick = f AND
referral_source = other AND
query_on_thyroxine = f AND
query_hyperthyroid = f AND
T3_measured = t AND
query_hypothyroid = t AND
T3 > 0.9: negative (7.0)

sick = f AND
FTI <= 66: sick (19.0/2.0)

sick = f AND
query_hypothyroid = f AND
T3_measured = t AND
TSH > 0.025 AND
referral_source = SVI AND
sex = F AND
FTI <= 176: negative (11.0/1.0)

sick = f AND
T3_measured = f: negative (8.0)

sick = f AND
query_hypothyroid = f AND
TSH > 0.025 AND
referral_source = SVI AND
sex = M: negative (6.0)

sick = f AND
query_hypothyroid = f AND
referral_source = other AND
TT4 <= 155 AND
T4U > 0.56 AND
TSH <= 34: negative (21.0/1.0)

TT4 <= 161: sick (21.0)

TSH <= 0.32 AND
referral_source = other AND
age > 54: negative (3.0/1.0)

TSH <= 0.32: sick (5.0)

: negative (3.0)

Number of Rules : 20

Time taken to build model: 0.12 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

==== Summary ===

    Correctly Classified Instances      1745          98.3653 %
    Incorrectly Classified Instances     29           1.6347 %
    Kappa statistic                      0.8539
    Mean absolute error                  0.0179
    Root mean squared error              0.1219
    Relative absolute error              15.2824 %
    Root relative squared error         51.1966 %
    Total Number of Instances            1774

==== Detailed Accuracy By Class ===

          TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
          0.992     0.150     0.990     0.992     0.991     0.854     0.970     0.996     negative
          0.850     0.008     0.875     0.850     0.863     0.854     0.970     0.807     sick
Weighted Avg.       0.984     0.141     0.983     0.984     0.984     0.854     0.970     0.985

==== Confusion Matrix ===

      a      b  <- classified as
1654   13 |   a = negative
      16   91 |   b = sick

```

## Percentage Split 60%

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3548
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 60.0% train, remainder test

==== Classifier model (full training set) ===

PART decision list
-----
T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3157.0/1.0)

referral_source = SVI AND
TSH <= 13 AND
FTI <= 157 AND
TT4 > 50 AND
FTI > 64: sick (157.0)

sick = f AND
T4U_measured = t AND
referral_source = other AND
age <= 77 AND
query_hypothyroid = f AND
T3_measured = t AND
TT4 <= 58: negative (31.0)

sick = f AND
T4U_measured = t AND
T3 > 2 AND
on_thyroxine = t: negative (24.0)

sick = f AND
T4U_measured = t AND
T3 <= 2.8 AND
TSH_measured = t AND
referral_source = SVI AND
TSH > 9.8: negative (20.0)

sick = f AND
T4U_measured = t AND
T3 > 2.8: negative (14.0)

sick = f AND
T4U_measured = f: negative (12.0)

sick = f AND
TSH_measured = f: negative (11.0/1.0)

sick = f AND
referral_source = SVHC: sick (9.0/1.0)

```

```

sick = f AND
referral_source = other AND
query_on_thyroxine = f AND
age > 77: sick (9.0/1.0)

sick = f AND
referral_source = other AND
query_on_thyroxine = f AND
query_hyperthyroid = f AND
T3_measured = t AND
query_hypothyroid = t AND
T3 > 0.9: negative (7.0)

sick = f AND
FTI <= 66: sick (19.0/2.0)

sick = f AND
query_hypothyroid = f AND
T3_measured = t AND
TSH > 0.025 AND
referral_source = SVI AND
sex = F AND
FTI <= 176: negative (11.0/1.0)

sick = f AND
T3_measured = f: negative (8.0)

sick = f AND
query_hypothyroid = f AND
TSH > 0.025 AND
referral_source = SVI AND
sex = M: negative (6.0)

sick = f AND
query_hypothyroid = f AND
referral_source = other AND
TT4 <= 155 AND
T4U > 0.56 AND
TSH <= 34: negative (21.0/1.0)

TT4 <= 161: sick (21.0)

TSH <= 0.32 AND
referral_source = other AND
age > 54: negative (3.0/1.0)

TSH <= 0.32: sick (5.0)

: negative (3.0)

Number of Rules : 20

Time taken to build model: 0.12 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

==== Summary ===

    Correctly Classified Instances      1393          98.1677 %
    Incorrectly Classified Instances     26           1.8323 %
    Kappa statistic                      0.8298
    Mean absolute error                  0.0186
    Root mean squared error              0.127
    Relative absolute error             16.2232 %
    Root relative squared error         56.6143 %
    Total Number of Instances           1419

==== Detailed Accuracy By Class ===

          TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
          0.986     0.093     0.995     0.986     0.990     0.832     0.970     0.997     negative
          0.907     0.014     0.782     0.907     0.840     0.832     0.970     0.866     sick
Weighted Avg.   0.982     0.089     0.983     0.982     0.982     0.832     0.970     0.990

==== Confusion Matrix ===

      a      b  <- classified as
1325   19 |   a = negative
      7   68 |   b = sick

```

## Percentage Split 66%

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3548
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 66.0% train, remainder test

==== Classifier model (full training set) ===

PART decision list
-----
T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3157.0/1.0)

referral_source = SVI AND
TSH <= 13 AND
FTI <= 157 AND
TT4 > 50 AND
FTI > 64: sick (157.0)

sick = f AND
T4U_measured = t AND
referral_source = other AND
age <= 77 AND
query_hypothyroid = f AND
T3_measured = t AND
TT4 <= 58: negative (31.0)

sick = f AND
T4U_measured = t AND
T3 > 2 AND
on_thyroxine = t: negative (24.0)

sick = f AND
T4U_measured = t AND
T3 <= 2.8 AND
TSH_measured = t AND
referral_source = SVI AND
TSH > 9.8: negative (20.0)

sick = f AND
T4U_measured = t AND
T3 > 2.8: negative (14.0)

sick = f AND
T4U_measured = f: negative (12.0)

sick = f AND
TSH_measured = f: negative (11.0/1.0)

sick = f AND
referral_source = SVHC: sick (9.0/1.0)

```

```

sick = f AND
referral_source = other AND
query_on_thyroxine = f AND
age > 77: sick (9.0/1.0)

sick = f AND
referral_source = other AND
query_on_thyroxine = f AND
query_hyperthyroid = f AND
T3_measured = t AND
query_hypothyroid = t AND
T3 > 0.9: negative (7.0)

sick = f AND
FTI <= 66: sick (19.0/2.0)

sick = f AND
query_hypothyroid = f AND
T3_measured = t AND
TSH > 0.025 AND
referral_source = SVI AND
sex = F AND
FTI <= 176: negative (11.0/1.0)

sick = f AND
T3_measured = f: negative (8.0)

sick = f AND
query_hypothyroid = f AND
TSH > 0.025 AND
referral_source = SVI AND
sex = M: negative (6.0)

sick = f AND
query_hypothyroid = f AND
referral_source = other AND
TT4 <= 155 AND
T4U > 0.56 AND
TSH <= 34: negative (21.0/1.0)

TT4 <= 161: sick (21.0)

TSH <= 0.32 AND
referral_source = other AND
age > 54: negative (3.0/1.0)

TSH <= 0.32: sick (5.0)

: negative (3.0)

Number of Rules : 20

Time taken to build model: 0.13 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

==== Summary ===

    Correctly Classified Instances      1186          98.3416 %
    Incorrectly Classified Instances     20           1.6584 %
    Kappa statistic                      0.835
    Mean absolute error                  0.0181
    Root mean squared error              0.1229
    Relative absolute error              16.1336 %
    Root relative squared error         55.8941 %
    Total Number of Instances            1206

==== Detailed Accuracy By Class ===

          TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
          0.989     0.115     0.994     0.989     0.991     0.836     0.955     0.995     negative
          0.885     0.011     0.806     0.885     0.844     0.836     0.955     0.816     sick
Weighted Avg.   0.983     0.110     0.984     0.983     0.984     0.836     0.955     0.986

==== Confusion Matrix ===

      a      b  <- classified as
1132    13 |    a = negative
      7    54 |    b = sick

```

## Leave one out fold

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3548
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   3548-fold cross-validation

==== Classifier model (full training set) ===

PART decision list
-----
T3 > 1.1 AND
FTI <= 171 AND
TT4 > 56: negative (3157.0/1.0)

referral_source = SVI AND
TSH <= 13 AND
FTI <= 157 AND
TT4 > 50 AND
FTI > 64: sick (157.0)

sick = f AND
T4U_measured = t AND
referral_source = other AND
age <= 77 AND
query_hypothyroid = f AND
T3_measured = t AND
TT4 <= 58: negative (31.0)

sick = f AND
T4U_measured = t AND
T3 > 2 AND
on_thyroxine = t: negative (24.0)

sick = f AND
T4U_measured = t AND
T3 <= 2.8 AND
TSH_measured = t AND
referral_source = SVI AND
TSH > 9.8: negative (20.0)

sick = f AND
T4U_measured = t AND
T3 > 2.8: negative (14.0)

sick = f AND
T4U_measured = f: negative (12.0)

sick = f AND
TSH_measured = f: negative (11.0/1.0)

sick = f AND
referral_source = SVHC: sick (9.0/1.0)

```

```

sick = f AND
referral_source = other AND
query_on_thyroxine = f AND
age > 77: sick (9.0/1.0)

sick = f AND
referral_source = other AND
query_on_thyroxine = f AND
query_hyperthyroid = f AND
T3_measured = t AND
query_hypothyroid = t AND
T3 > 0.9: negative (7.0)

sick = f AND
FTI <= 66: sick (19.0/2.0)

sick = f AND
query_hypothyroid = f AND
T3_measured = t AND
TSH > 0.025 AND
referral_source = SVI AND
sex = F AND
FTI <= 176: negative (11.0/1.0)

sick = f AND
T3_measured = f: negative (8.0)

sick = f AND
query_hypothyroid = f AND
TSH > 0.025 AND
referral_source = SVI AND
sex = M: negative (6.0)

sick = f AND
query_hypothyroid = f AND
referral_source = other AND
TT4 <= 155 AND
T4U > 0.56 AND
TSH <= 34: negative (21.0/1.0)

TT4 <= 161: sick (21.0)

TSH <= 0.32 AND
referral_source = other AND
age > 54: negative (3.0/1.0)

TSH <= 0.32: sick (5.0)

: negative (3.0)

Number of Rules : 20

Time taken to build model: 0.14 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3499          98.6189 %
Incorrectly Classified Instances   49           1.3811 %
Kappa statistic                   0.8815
Mean absolute error               0.0147
Root mean squared error          0.113
Relative absolute error           12.5756 %
Root relative squared error     46.7517 %
Total Number of Instances        3548

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall    F-Measure   MCC      ROC Area   PRC Area   Class
      0.993     0.113     0.992     0.993     0.993     0.882     0.935     0.992   negative
      0.887     0.007     0.891     0.887     0.889     0.882     0.935     0.870     sick
Weighted Avg.      0.986     0.107     0.986     0.986     0.986     0.882     0.935     0.985

==== Confusion Matrix ====

      a      b  <- classified as
3303   24 |      a = negative
      25   196 |      b = sick

```

Preprocessing method	PART					
	Cross Validation			Percentage Split		
	5 folds	10 folds	15 folds	50.00%	60.00%	66.00%
With Original Dataset	98.3033%	98.6214%	98.5949%	98.6214%	97.8131%	98.7520%
Missing Values only	98.1177%	98.7805%	98.7010%	96.9777%	97.8131%	98.9860%
Missing Values + Extreme Values	98.2079%	98.3457%	98.3182%	98.2350%	98.3460%	98.2157%
Missing Values + Outliers	98.3089%	98.5626%	98.5062%	98.3653%	98.1677%	98.3416%

### 3.5.4 Conclusion

Dealing with missing values yields the best results of all other preprocessing tests. Using Missing Values filter gave 3 best results in Cross Validation 10 and 15 folds. and in Percentage split 66%. dealing with extreme values won in Percentage split 60% while dealing with outliers won Cross validation 5 folds. which indicates that PART can benefit from dealing with missing values.

## 3.6 Combos

### 3.6.1 Missing Values + Outliers + Extreme Values

#### Cross Validation 5 folds

```
==== Run information ====
Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3439
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   5-fold cross-validation

==== Classifier model (full training set) ====
PART decision list
-----
T3 > 1.1 AND
TT4 > 57 AND
FTI <= 163: negative (3066.0)

FTI > 155 AND
on_thyroxine = t: negative (43.0)

T3 <= 1.1 AND
referral_source = SVI AND
FTI <= 153 AND
```

```

TT4 > 49 AND
FTI > 69: sick (153.0)

sick = f AND
T4U_measured = t AND
T3 > 2.7: negative (15.0)

T4U_measured = t AND
sick = f AND
TSH_measured = t AND
FTI <= 52: sick (12.0/1.0)

T3_measured = f: negative (20.0)

sick = f AND
T4U_measured = f: negative (8.0)

sick = f AND
query_hypothyroid = t AND
TT4 <= 156: sick (5.0)

sick = f AND
query_hypothyroid = f AND
referral_source = other AND
TT4 > 155 AND
T3 > 1.7: sick (7.0)

sick = f AND
referral_source = other AND
age <= 70: negative (37.0/2.0)

sick = f AND
FTI > 157 AND
FTI <= 176: negative (20.0/1.0)

sick = f AND
age > 77: sick (11.0)

sick = f AND
referral_source = SVHC AND
T3 <= 1.5: sick (5.0)

sick = f AND
referral_source = SVI AND
TSH > 0.3 AND
TT4 > 49 AND
TT4 <= 83: sick (8.0/1.0)

sick = f AND
TSH > 0.3 AND
referral_source = SVI: negative (9.0)

sick = f AND
referral_source = SVI: sick (6.0)

sick = f AND
T3 > 0.8: negative (5.0)

: sick (9.0/1.0)

Number of Rules : 18

Time taken to build model: 0.12 seconds

==== Stratified cross-validation ====
==== Summary ====

  Correctly Classified Instances      3393          98.6624 %
  Incorrectly Classified Instances     46           1.3376 %
  Kappa statistic                      0.8828
  Mean absolute error                  0.0144
  Root mean squared error              0.1083
  Relative absolute error              12.2165 %
  Root relative squared error         44.6324 %
  Total Number of Instances            3439

==== Detailed Accuracy By Class ====

    TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
    0.995     0.139     0.991      0.995     0.993     0.883     0.964     0.995     negative
    0.861     0.005     0.921      0.861     0.890     0.883     0.964     0.870     sick
  Weighted Avg.     0.987     0.130     0.986      0.987     0.986     0.883     0.964     0.987

==== Confusion Matrix ====

  a      b  <-- classified as

```

```
3207 16 | a = negative
 30 186 | b = sick
```

## Cross Validation 10 folds

```
==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3439
Attributes:  30
             age
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             I131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary
             psych
             TSH_measured
             TSH
             T3_measured
             T3
             TT4_measured
             TT4
             T4U_measured
             T4U
             FTI_measured
             FTI
             TBG_measured
             TBG
             referral_source
             Class
Test mode:   10-fold cross-validation

==== Classifier model (full training set) ===

PART decision list
-----
T3 > 1.1 AND
TT4 > 57 AND
FTI <= 163: negative (3066.0)

FTI > 155 AND
on_thyroxine = t: negative (43.0)

T3 <= 1.1 AND
referral_source = SVI AND
FTI <= 153 AND
TT4 > 49 AND
FTI > 69: sick (153.0)

sick = f AND
T4U_measured = t AND
T3 > 2.7: negative (15.0)

T4U_measured = t AND
sick = f AND
TSH_measured = t AND
FTI <= 52: sick (12.0/1.0)

T3_measured = f: negative (20.0)

sick = f AND
T4U_measured = f: negative (8.0)

sick = f AND
query_hypothyroid = f AND
referral_source = other AND
TT4 > 155 AND
T3 > 1.7: sick (7.0)
```

```

sick = f AND
referral_source = other AND
age <= 70: negative (37.0/2.0)

sick = f AND
FTI > 157 AND
FTI <= 176: negative (20.0/1.0)

sick = f AND
age > 77: sick (11.0)

sick = f AND
referral_source = SVHC AND
T3 <= 1.5: sick (5.0)

sick = f AND
referral_source = SVI AND
TSH > 0.3 AND
TT4 > 49 AND
TT4 <= 83: sick (8.0/1.0)

sick = f AND
TSH > 0.3 AND
referral_source = SVI: negative (9.0)

sick = f AND
referral_source = SVI: sick (6.0)

sick = f AND
T3 > 0.8: negative (5.0)

: sick (9.0/1.0)

Number of Rules : 18

Time taken to build model: 0.07 seconds

==== Stratified cross-validation ====
==== Summary ====

    Correctly Classified Instances      3387          98.4879 %
    Incorrectly Classified Instances   52           1.5121 %
    Kappa statistic                   0.871
    Mean absolute error              0.0155
    Root mean squared error          0.1147
    Relative absolute error          13.1538 %
    Root relative squared error     47.2655 %
    Total Number of Instances        3439

==== Detailed Accuracy By Class ====

    TP Rate   FP Rate   Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
    0.992     0.125     0.992     0.992    0.992     0.871    0.978    0.997    negative
    0.875     0.008     0.883     0.875    0.879     0.871    0.978    0.883    sick
    Weighted Avg. 0.985     0.118     0.985     0.985    0.985     0.871    0.978    0.990

==== Confusion Matrix ====

    a      b  <- classified as
3198   25 |  a = negative
    27   189 |  b = sick

```

## Cross Validation 15 folds

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3439
Attributes:  30
             age
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             I131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium

```

```

goitre
tumor
hypopituitary
psych
TSH_measured
TSH
T3_measured
T3
TT4_measured
TT4
T4U_measured
T4U
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Test mode: 15-fold cross-validation

==== Classifier model (full training set) ====

PART decision list
-----
T3 > 1.1 AND
TT4 > 57 AND
FTI <= 163: negative (3066.0)

FTI > 155 AND
on_thyroxine = t: negative (43.0)

T3 <= 1.1 AND
referral_source = SVI AND
FTI <= 153 AND
TT4 > 49 AND
FTI > 69: sick (153.0)

sick = f AND
T4U_measured = t AND
T3 > 2.7: negative (15.0)

T4U_measured = t AND
sick = f AND
TSH_measured = t AND
FTI <= 52: sick (12.0/1.0)

T3_measured = f: negative (20.0)

sick = f AND
T4U_measured = f: negative (8.0)

sick = f AND
query_hypothyroid = t AND
TT4 <= 156: sick (5.0)

sick = f AND
query_hypothyroid = f AND
referral_source = other AND
TT4 > 155 AND
T3 > 1.7: sick (7.0)

sick = f AND
referral_source = other AND
age <= 70: negative (37.0/2.0)

sick = f AND
FTI > 157 AND
FTI <= 176: negative (20.0/1.0)

sick = f AND
age > 77: sick (11.0)

sick = f AND
referral_source = SVHC AND
T3 <= 1.5: sick (5.0)

sick = f AND
referral_source = SVI AND
TSH > 0.3 AND
TT4 > 49 AND
TT4 <= 83: sick (8.0/1.0)

sick = f AND
TSH > 0.3 AND
referral_source = SVI: negative (9.0)

```

```

sick = f AND
referral_source = SVI: sick (6.0)

sick = f AND
T3 > 0.8: negative (5.0)

: sick (9.0/1.0)

Number of Rules : 18

Time taken to build model: 0.08 seconds

==== Stratified cross-validation ====
==== Summary ===

Correctly Classified Instances      3401          98.895 %
Incorrectly Classified Instances   38           1.105 %
Kappa statistic                   0.9069
Mean absolute error               0.012
Root mean squared error          0.1011
Relative absolute error          10.1977 %
Root relative squared error     41.6546 %
Total Number of Instances        3439

==== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC    ROC Area  PRC Area  Class
0.993     0.079    0.995     0.993    0.994     0.907    0.970    0.996    negative
0.921     0.007    0.905     0.921    0.913     0.907    0.970    0.886    sick
Weighted Avg. 0.989     0.074    0.989     0.989    0.989     0.907    0.970    0.989

==== Confusion Matrix ====

      a     b  <-- classified as
3202   21 |     a = negative
      17  199 |     b = sick

```

## Percentage Split 50%

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3439
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 50.0% train, remainder test

==== Classifier model (full training set) ===

PART decision list
-----
T3 > 1.1 AND

```

```

TT4 > 57 AND
FTI <= 163: negative (3066.0)

FTI > 155 AND
on_thyroxine = t: negative (43.0)

T3 <= 1.1 AND
referral_source = SVI AND
FTI <= 153 AND
TT4 > 49 AND
FTI > 69: sick (153.0)

sick = f AND
T4U_measured = t AND
T3 > 2.7: negative (15.0)

T4U_measured = t AND
sick = f AND
TSH_measured = t AND
FTI <= 52: sick (12.0/1.0)

T3_measured = f: negative (20.0)

sick = f AND
T4U_measured = f: negative (8.0)

sick = f AND
query_hypothyroid = t AND
TT4 <= 156: sick (5.0)

sick = f AND
query_hypothyroid = f AND
referral_source = other AND
TT4 > 155 AND
T3 > 1.7: sick (7.0)

sick = f AND
referral_source = other AND
age <= 70: negative (37.0/2.0)

sick = f AND
FTI > 157 AND
FTI <= 176: negative (20.0/1.0)

sick = f AND
age > 77: sick (11.0)

sick = f AND
referral_source = SVHC AND
T3 <= 1.5: sick (5.0)

sick = f AND
referral_source = SVI AND
TSH > 0.3 AND
TT4 > 49 AND
TT4 <= 83: sick (8.0/1.0)

sick = f AND
TSH > 0.3 AND
referral_source = SVI: negative (9.0)

sick = f AND
referral_source = SVI: sick (6.0)

sick = f AND
T3 > 0.8: negative (5.0)

: sick (9.0/1.0)

Number of Rules : 18

Time taken to build model: 0.1 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0.01 seconds

==== Summary ===

Correctly Classified Instances      1695           98.6038 %
Incorrectly Classified Instances     24            1.3962 %
Kappa statistic                      0.8676
Mean absolute error                  0.0161
Root mean squared error              0.1141
Relative absolute error              13.6186 %
Root relative squared error         48.929  %

```

```
Total Number of Instances          1719
==== Detailed Accuracy By Class ====
      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.994     0.152     0.991     0.994     0.993     0.868     0.946     0.993     negative
      0.848     0.006     0.903     0.848     0.875     0.868     0.946     0.806     sick
Weighted Avg.    0.986     0.143     0.986     0.986     0.986     0.868     0.946     0.983
==== Confusion Matrix ====
      a      b  <-- classified as
1611    9 |      a = negative
      15    84 |      b = sick
```

## Percentage Split 60%

```
==== Run information ====
Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3439
Attributes:  30
            age
            sex
            on_thyroxine
            query_on_thyroxine
            on_antithyroid_medication
            sick
            pregnant
            thyroid_surgery
            I131_treatment
            query_hypothyroid
            query_hyperthyroid
            lithium
            goitre
            tumor
            hypopituitary
            psych
            TSH_measured
            TSH
            T3_measured
            T3
            TT4_measured
            TT4
            T4U_measured
            T4U
            FTI_measured
            FTI
            TBG_measured
            TBG
            referral_source
            Class
Test mode:   split 60.0% train, remainder test
==== Classifier model (full training set) ====
PART decision list
-----
T3 > 1.1 AND
TT4 > 57 AND
FTI <= 163: negative (3066.0)

FTI > 155 AND
on_thyroxine = t: negative (43.0)

T3 <= 1.1 AND
referral_source = SVI AND
FTI <= 153 AND
TT4 > 49 AND
FTI > 69: sick (153.0)

sick = f AND
T4U_measured = t AND
T3 > 2.7: negative (15.0)

T4U_measured = t AND
sick = f AND
TSH_measured = t AND
FTI <= 52: sick (12.0/1.0)

T3_measured = f: negative (20.0)
```

```

sick = f AND
T4U_measured = f: negative (8.0)

sick = f AND
query_hypothyroid = t AND
TT4 <= 156: sick (5.0)

sick = f AND
query_hypothyroid = f AND
referral_source = other AND
TT4 > 155 AND
T3 > 1.7: sick (7.0)

sick = f AND
referral_source = other AND
age <= 70: negative (37.0/2.0)

sick = f AND
FTI > 157 AND
FTI <= 176: negative (20.0/1.0)

sick = f AND
age > 77: sick (11.0)

sick = f AND
referral_source = SVHC AND
T3 <= 1.5: sick (5.0)

sick = f AND
referral_source = SVI AND
TSH > 0.3 AND
TT4 > 49 AND
TT4 <= 83: sick (8.0/1.0)

sick = f AND
TSH > 0.3 AND
referral_source = SVI: negative (9.0)

sick = f AND
referral_source = SVI: sick (6.0)

sick = f AND
T3 > 0.8: negative (5.0)

: sick (9.0/1.0)

Number of Rules : 18

Time taken to build model: 0.08 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0.01 seconds

==== Summary ===

  Correctly Classified Instances      1353          98.3285 %
  Incorrectly Classified Instances     23           1.6715 %
  Kappa statistic                      0.855
  Mean absolute error                  0.0187
  Root mean squared error              0.1261
  Relative absolute error             15.7293 %
  Root relative squared error         51.0069 %
  Total Number of Instances            1376

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
  0.995     0.189     0.987     0.995     0.991     0.857     0.867     0.984   negative
  0.811     0.005     0.924     0.811     0.864     0.857     0.867     0.813   sick
Weighted Avg.    0.983     0.177     0.983     0.983     0.983     0.857     0.867     0.972

==== Confusion Matrix ===

    a      b  <- classified as
1280     6 |     a = negative
    17    73 |     b = sick

```

## Percentage Split 66%

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2

```

```

Relation: sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances: 3439
Attributes: 30
age
sex
on_thyroxine
query_on_thyroxine
on_antithyroid_medication
sick
pregnant
thyroid_surgery
I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH
T3_measured
T3
TT4_measured
TT4
T4U_measured
T4U
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Test mode: split 66.0% train, remainder test

==== Classifier model (full training set) ====

PART decision list
-----
T3 > 1.1 AND
TT4 > 57 AND
FTI <= 163: negative (3066.0)

FTI > 155 AND
on_thyroxine = t: negative (43.0)

T3 <= 1.1 AND
referral_source = SVI AND
FTI <= 153 AND
TT4 > 49 AND
FTI > 69: sick (153.0)

sick = f AND
T4U_measured = t AND
T3 > 2.7: negative (15.0)

T4U_measured = t AND
sick = f AND
TSH_measured = t AND
FTI <= 52: sick (12.0/1.0)

T3_measured = f: negative (20.0)

sick = f AND
T4U_measured = f: negative (8.0)

sick = f AND
query_hypothyroid = t AND
TT4 <= 156: sick (5.0)

sick = f AND
query_hypothyroid = f AND
referral_source = other AND
TT4 > 155 AND
T3 > 1.7: sick (7.0)

sick = f AND
referral_source = other AND
age <= 70: negative (37.0/2.0)

sick = f AND
FTI > 157 AND
FTI <= 176: negative (20.0/1.0)

sick = f AND
age > 77: sick (11.0)

```

```

sick = f AND
referral_source = SVHC AND
T3 <= 1.5: sick (5.0)

sick = f AND
referral_source = SVI AND
TSH > 0.3 AND
TT4 > 49 AND
TT4 <= 83: sick (8.0/1.0)

sick = f AND
TSH > 0.3 AND
referral_source = SVI: negative (9.0)

sick = f AND
referral_source = SVI: sick (6.0)

sick = f AND
T3 > 0.8: negative (5.0)

: sick (9.0/1.0)

Number of Rules : 18

Time taken to build model: 0.17 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0.01 seconds

==== Summary ===

Correctly Classified Instances      1142          97.6903 %
Incorrectly Classified Instances   27           2.3097 %
Kappa statistic                   0.7785
Mean absolute error               0.0242
Root mean squared error          0.1411
Relative absolute error          20.7945 %
Root relative squared error     59.8486 %
Total Number of Instances        1169

==== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC    ROC Area  PRC Area  Class
      0.992    0.261    0.984    0.992    0.988    0.781    0.937    0.992    negative
      0.739    0.008    0.850    0.739    0.791    0.781    0.937    0.782    sick
Weighted Avg.      0.977    0.246    0.976    0.977    0.976    0.781    0.937    0.980

==== Confusion Matrix ===

      a      b  <- classified as
1091    9 |  a = negative
      18    51 |  b = sick

```

## Leave one out fold

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3439
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3

```

```

TT4_measured
TT4
T4U_measured
T4U
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Test mode: 3439-fold cross-validation

==== Classifier model (full training set) ====

PART decision list
-----
T3 > 1.1 AND
TT4 > 57 AND
FTI <= 163: negative (3066.0)

FTI > 155 AND
on_thyroxine = t: negative (43.0)

T3 <= 1.1 AND
referral_source = SVI AND
FTI <= 153 AND
TT4 > 49 AND
FTI > 69: sick (153.0)

sick = f AND
T4U_measured = t AND
T3 > 2.7: negative (15.0)

T4U_measured = t AND
sick = f AND
TSH_measured = t AND
FTI <= 52: sick (12.0/1.0)

T3_measured = f: negative (20.0)

sick = f AND
T4U_measured = f: negative (8.0)

sick = f AND
query_hypothyroid = t AND
referral_source = other AND
TT4 <= 156: sick (5.0)

sick = f AND
query_hypothyroid = f AND
referral_source = other AND
TT4 > 155 AND
T3 > 1.7: sick (7.0)

sick = f AND
referral_source = other AND
age <= 70: negative (37.0/2.0)

sick = f AND
FTI > 157 AND
FTI <= 176: negative (20.0/1.0)

sick = f AND
age > 77: sick (11.0)

sick = f AND
referral_source = SVHC AND
T3 <= 1.5: sick (5.0)

sick = f AND
referral_source = SVI AND
TSH > 0.3 AND
TT4 > 49 AND
TT4 <= 83: sick (8.0/1.0)

sick = f AND
TSH > 0.3 AND
referral_source = SVI: negative (9.0)

sick = f AND
referral_source = SVI: sick (6.0)

sick = f AND
T3 > 0.8: negative (5.0)

: sick (9.0/1.0)

```

```

Number of Rules : 18

Time taken to build model: 0.09 seconds

==== Stratified cross-validation ====
==== Summary ===

    Correctly Classified Instances      3398          98.8078 %
    Incorrectly Classified Instances   41           1.1922 %
    Kappa statistic                   0.9002
    Mean absolute error              0.0133
    Root mean squared error          0.104
    Relative absolute error          11.2786 %
    Root relative squared error     42.8593 %
    Total Number of Instances        3439

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC   ROC Area   PRC Area   Class
      0.993     0.079     0.995     0.993     0.994     0.900     0.983     0.998     negative
      0.921     0.007     0.892     0.921     0.907     0.900     0.983     0.917     sick
Weighted Avg.   0.988     0.074     0.988     0.988     0.988     0.900     0.983     0.993

==== Confusion Matrix ====

      a     b  <-- classified as
3199  24 |     a = negative
      17 199 |     b = sick

```

### 3.6.2 Missing + Extreme + Outliers + frequency 40

#### Cross Validation 5 folds

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3439
Attributes:  248
             [list of attributes omitted]
Test mode:   5-fold cross-validation

==== Classifier model (full training set) ===

PART decision list
-----
T3_3 = '(1.15-inf)' AND
FTI_39 = '(-inf-170.5)' AND
TT4_1 = '(61.5-inf)': negative (3079.0/1.0)

T3_3 = '(1.15-inf)' AND
T4U_33 = '(-inf-1.115)' AND
referral_source = other AND
TSH_12 = '(0.765-inf)': negative (27.0)

T3_19 = '(-inf-2.65)' AND
T3_9 = '(-inf-1.75)' AND
TT4_39 = '(-inf-169.5)' AND
FTI_38 = '(-inf-157.5)' AND
TT4_37 = '(-inf-147.5)' AND
referral_source = SVI AND
TT4_1 = '(61.5-inf)': sick (146.0/1.0)

T3_19 = '(2.65-inf)': negative (17.0)

T4U_measured = t AND
age_37 = '(77.5-inf)' AND
TT4_32 = '(-inf-127.5)': sick (15.0)

T4U_measured = t AND
TT4_17 = '(101.5-inf)' AND
on_thyroxine = t: negative (24.0)

goitre = f AND
T4U_33 = '(-inf-1.115)' AND
T4U_measured = t AND
on_thyroxine = f AND
age_1 = '(17.5-inf)' AND
T3_measured = t AND
TSH_2 = '(0.0375-inf)' AND
sick = f AND

```

```

TSH_37 = '(-inf-6.75]' AND
T3_1 = '(0.75-inf)' AND
T4U_2 = '(0.745-inf)' AND
T4U_31 = '(-inf-1.075]' AND
T4U_27 = '(-inf-1.035]' AND
TSH_7 = '(0.295-inf)' AND
FTI_16 = '(104.5-inf)': negative (13.0)

```

```

T4U_measured = t AND
on_thyroxine = f AND
age_1 = '(17.5-inf)' AND
goitre = f AND
T3_17 = '(-inf-2.45]' AND
T4U_33 = '(-inf-1.115]' AND
T3_measured = t AND
TSH_2 = '(0.0375-inf)' AND
sick = f AND
age_3 = '(22.5-inf)' AND
T3_12 = '(-inf-2.00675]' AND
TSH_3 = '(0.075-inf)' AND
TSH_37 = '(-inf-6.75]' AND
age_34 = '(-inf-72.5]' AND
T3_9 = '(-inf-1.75]' AND
age_9 = '(34.5-inf)' AND
referral_source = other: negative (17.0/1.0)

```

```

T4U_measured = t AND
age_1 = '(17.5-inf)' AND
on_thyroxine = f AND
T4U_29 = '(-inf-1.055]' AND
T3_measured = f AND
sex = F: negative (9.0)

```

```

T4U_measured = t AND
age_1 = '(17.5-inf)' AND
TT4_38 = '(-inf-157.5]' AND
age_37 = '(-inf-77.51]' AND
T3_13 = '(-inf-2.05675]' AND
TSH_3 = '(0.075-inf)' AND
TT4_1 = '(61.5-inf)': sick (19.0/1.0)

```

```

TT4_38 = '(-inf-157.5]' AND
T4U_28 = '(-inf-1.045]' AND
goitre = f AND
T4U_13 = '(0.905-inf)': negative (14.0)

```

```

age_1 = '(17.5-inf)' AND
psych = f AND
T4U_11 = '(0.885-inf)': sick (13.0)

```

```

age_5 = '(26.5-inf)' AND
psych = f AND
TT4_38 = '(-inf-157.5]' AND
age_37 = '(-inf-77.51]' AND
T3_3 = '(1.15-inf)' AND
age_16 = '(-inf-48.5)': negative (5.0)

```

```

age_5 = '(26.5-inf)' AND
psych = f AND
age_15 = '(46.5-inf)' AND
TT4_38 = '(-inf-157.5]' AND
age_37 = '(-inf-77.51]' AND
TSH_14 = '(0.925-inf)' AND
TSH_28 = '(-inf-2.55)': sick (7.0)

```

```

TT4_38 = '(-inf-157.5]' AND
age_5 = '(26.5-inf)' AND
age_34 = '(-inf-72.5]' AND
psych = f AND
age_20 = '(-inf-55.5)': sick (10.0/1.0)

```

```

TT4_38 = '(-inf-157.5]' AND
age_28 = '(-inf-63.5)': negative (9.0)

```

```

TT4_38 = '(-inf-157.5]' AND
TSH_12 = '(-inf-0.765)': negative (9.0/1.0)

```

```
: sick (6.0)
```

```
Number of Rules : 18
```

```
Time taken to build model: 0.32 seconds
```

```
==== Stratified cross-validation ====
==== Summary ====

```

```

Correctly Classified Instances      3378          98.2262 %
Incorrectly Classified Instances   61           1.7738 %
Kappa statistic                   0.8484
Mean absolute error               0.0188
Root mean squared error          0.1283
Relative absolute error           15.944 %
Root relative squared error     52.8844 %
Total Number of Instances        3439

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC    ROC Area   PRC Area   Class
          0.991     0.148     0.990     0.991     0.991     0.848     0.902     0.984   negative
          0.852     0.009     0.864     0.852     0.858     0.848     0.902     0.798   sick
Weighted Avg.       0.982     0.139     0.982     0.982     0.982     0.848     0.902     0.972

==== Confusion Matrix ====

      a     b  <-- classified as
3194  29 |     a = negative
      32  184 |     b = sick

```

## Cross Validation 10 folds

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3439
Attributes:  248
              [list of attributes omitted]
Test mode:   10-fold cross-validation

==== Classifier model (full training set) ===

PART decision list
-----
T3_3 = '(1.15-inf)' AND
FTI_39 = '(-inf-170.5)' AND
TT4_1 = '(61.5-inf)': negative (3079.0/1.0)

T3_3 = '(1.15-inf)' AND
T4U_33 = '(-inf-1.115)' AND
referral_source = other AND
TSH_12 = '(0.765-inf)': negative (27.0)

T3_19 = '(-inf-2.65]' AND
T3_9 = '(-inf-1.75]' AND
TT4_39 = '(-inf-169.5)' AND
FTI_38 = '(-inf-157.5)' AND
TT4_37 = '(-inf-147.5)' AND
referral_source = SVI AND
TT4_1 = '(61.5-inf)': sick (146.0/1.0)

T3_19 = '(2.65-inf)': negative (17.0)

T4U_measured = t AND
age_37 = '(77.5-inf)' AND
TT4_32 = '(-inf-127.5)': sick (15.0)

T4U_measured = t AND
TT4_17 = '(101.5-inf)' AND
on_thyroxine = t: negative (24.0)

goitre = f AND
T4U_33 = '(-inf-1.115]' AND
T4U_measured = t AND
on_thyroxine = f AND
age_1 = '(17.5-inf)' AND
T3_measured = t AND
TSH_2 = '(0.0375-inf)' AND
sick = f AND
TSH_37 = '(-inf-6.75]' AND
T3_1 = '(0.75-inf)' AND
T4U_2 = '(0.745-inf)' AND
T4U_31 = '(-inf-1.075]' AND
T4U_27 = '(-inf-1.035)' AND
TSH_7 = '(0.295-inf)' AND
FTI_16 = '(104.5-inf)': negative (13.0)

T4U_measured = t AND
on_thyroxine = f AND
age_1 = '(17.5-inf)' AND

```

```

goitre = f AND
T3_17 = '(-inf-2.45]' AND
T4U_33 = '(-inf-1.115]' AND
T3_measured = t AND
TSH_2 = '(0.0375-inf)' AND
sick = f AND
age_3 = '(22.5-inf)' AND
T3_12 = '(-inf-2.00675]' AND
TSH_3 = '(0.075-inf)' AND
TSH_37 = '(-inf-6.75]' AND
age_34 = '(-inf-72.5]' AND
T3_9 = '(-inf-1.75]' AND
age_9 = '(34.5-inf)' AND
referral_source = other: negative (17.0/1.0)

T4U_measured = t AND
age_1 = '(17.5-inf)' AND
on_thyroxine = f AND
T4U_29 = '(-inf-1.055]' AND
T3_measured = f AND
sex = F: negative (9.0)

T4U_measured = t AND
age_1 = '(17.5-inf)' AND
TT4_38 = '(-inf-157.5]' AND
age_37 = '(-inf-77.5]' AND
T3_13 = '(-inf-2.05675]' AND
TSH_3 = '(0.075-inf)' AND
TT4_1 = '(61.5-inf)': sick (19.0/1.0)

TT4_38 = '(-inf-157.5]' AND
T4U_28 = '(-inf-1.045]' AND
goitre = f AND
T4U_13 = '(0.905-inf)': negative (14.0)

age_1 = '(17.5-inf)' AND
psych = f AND
T4U_11 = '(0.885-inf)': sick (13.0)

age_5 = '(26.5-inf)' AND
psych = f AND
TT4_38 = '(-inf-157.5]' AND
age_37 = '(-inf-77.5]' AND
T3_3 = '(1.15-inf)' AND
age_16 = '(-inf-48.5)': negative (5.0)

age_5 = '(26.5-inf)' AND
psych = f AND
age_15 = '(46.5-inf)' AND
TT4_38 = '(-inf-157.5]' AND
age_37 = '(-inf-77.5]' AND
TSH_14 = '(0.925-inf)' AND
TSH_28 = '(-inf-2.55)': sick (7.0)

TT4_38 = '(-inf-157.5]' AND
age_5 = '(26.5-inf)' AND
age_34 = '(-inf-72.5]' AND
psych = f AND
age_20 = '(-inf-55.5)': sick (10.0/1.0)

TT4_38 = '(-inf-157.5]' AND
age_28 = '(-inf-63.5)': negative (9.0)

TT4_38 = '(-inf-157.5]' AND
TSH_12 = '(-inf-0.765)': negative (9.0/1.0)

: sick (6.0)

Number of Rules : 18

Time taken to build model: 0.33 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3387           98.4879 %
Incorrectly Classified Instances    52            1.5121 %
Kappa statistic                   0.8699
Mean absolute error               0.0161
Root mean squared error           0.1199
Relative absolute error            13.6561 %
Root relative squared error       49.4179 %
Total Number of Instances         3439

==== Detailed Accuracy By Class ====

```

```

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
      0.993    0.134    0.991     0.993    0.992     0.870    0.913     0.985    negative
      0.866    0.007    0.890     0.866    0.878     0.870    0.913     0.784    sick
Weighted Avg.    0.985    0.126    0.985     0.985    0.985     0.870    0.913     0.972

==== Confusion Matrix ====
a      b  <-- classified as
3200  23 |  a = negative
 29  187 |  b = sick

```

## Cross Validation 15 folds

```

==== Run information ====
Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3439
Attributes:  248
              [list of attributes omitted]
Test mode:   15-fold cross-validation

==== Classifier model (full training set) ====
PART decision list
-----
T3_3 = '(1.15-inf)' AND
FTI_39 = '(-inf-170.5)' AND
TT4_1 = '(61.5-inf)': negative (3079.0/1.0)

T3_3 = '(1.15-inf)' AND
T4U_33 = '(-inf-1.115)' AND
referral_source = other AND
TSH_12 = '(0.765-inf)': negative (27.0)

T3_19 = '(-inf-2.65)' AND
T3_9 = '(-inf-1.75)' AND
TT4_39 = '(-inf-169.5)' AND
FTI_38 = '(-inf-157.5)' AND
TT4_37 = '(-inf-147.5)' AND
referral_source = SVI AND
TT4_1 = '(61.5-inf)': sick (146.0/1.0)

T3_19 = '(2.65-inf)': negative (17.0)

T4U_measured = t AND
age_37 = '(77.5-inf)' AND
TT4_32 = '(-inf-127.5)': sick (15.0)

T4U_measured = t AND
TT4_17 = '(101.5-inf)' AND
on_thyroxine = t: negative (24.0)

goitre = f AND
T4U_33 = '(-inf-1.115)' AND
T4U_measured = t AND
on_thyroxine = f AND
age_1 = '(17.5-inf)' AND
T3_measured = t AND
TSH_2 = '(0.0375-inf)' AND
sick = f AND
TSH_37 = '(-inf-6.75)' AND
T3_1 = '(0.75-inf)' AND
T4U_2 = '(0.745-inf)' AND
T4U_31 = '(-inf-1.075)' AND
T4U_27 = '(-inf-1.035)' AND
TSH_7 = '(0.295-inf)' AND
FTI_16 = '(104.5-inf)': negative (13.0)

T4U_measured = t AND
on_thyroxine = f AND
age_1 = '(17.5-inf)' AND
goitre = f AND
T3_17 = '(-inf-2.45)' AND
T4U_33 = '(-inf-1.115)' AND
T3_measured = t AND
TSH_2 = '(0.0375-inf)' AND
sick = f AND
age_3 = '(22.5-inf)' AND
T3_12 = '(-inf-2.00675)' AND
TSH_3 = '(0.075-inf)' AND
TSH_37 = '(-inf-6.75)' AND
age_34 = '(-inf-72.5)' AND

```

```

T3_9 = '(-inf-1.75]' AND
age_9 = '(34.5-inf)' AND
referral_source = other: negative (17.0/1.0)

T4U_measured = t AND
age_1 = '(17.5-inf)' AND
on_thyroxine = f AND
T4U_29 = '(-inf-1.055]' AND
T3_measured = f AND
sex = F: negative (9.0)

T4U_measured = t AND
age_1 = '(17.5-inf)' AND
TT4_38 = '(-inf-157.5]' AND
age_37 = '(-inf-77.5]' AND
T3_13 = '(-inf-2.05675]' AND
TSH_3 = '(0.075-inf)' AND
TT4_1 = '(61.5-inf)': sick (19.0/1.0)

TT4_38 = '(-inf-157.5]' AND
T4U_28 = '(-inf-1.045]' AND
goitre = f AND
T4U_13 = '(0.905-inf)': negative (14.0)

age_1 = '(17.5-inf)' AND
psych = f AND
T4U_11 = '(0.885-inf)': sick (13.0)

age_5 = '(26.5-inf)' AND
psych = f AND
TT4_38 = '(-inf-157.5]' AND
age_37 = '(-inf-77.5]' AND
T3_3 = '(1.15-inf)' AND
age_16 = '(-inf-48.5)': negative (5.0)

age_5 = '(26.5-inf)' AND
psych = f AND
age_15 = '(46.5-inf)' AND
TT4_38 = '(-inf-157.5]' AND
age_37 = '(-inf-77.5]' AND
TSH_14 = '(0.925-inf)' AND
TSH_28 = '(-inf-2.55)': sick (7.0)

TT4_38 = '(-inf-157.5]' AND
age_5 = '(26.5-inf)' AND
age_34 = '(-inf-72.5]' AND
psych = f AND
age_20 = '(-inf-55.5)': sick (10.0/1.0)

TT4_38 = '(-inf-157.5]' AND
age_28 = '(-inf-63.5)': negative (9.0)

TT4_38 = '(-inf-157.5]' AND
TSH_12 = '(-inf-0.765)': negative (9.0/1.0)

: sick (6.0)

Number of Rules : 18

Time taken to build model: 0.31 seconds

==== Stratified cross-validation ====
==== Summary ====

    Correctly Classified Instances      3394          98.6915 %
    Incorrectly Classified Instances     45           1.3085 %
    Kappa statistic                   0.8876
    Mean absolute error               0.0143
    Root mean squared error          0.1106
    Relative absolute error          12.1144 %
    Root relative squared error     45.5898 %
    Total Number of Instances        3439

==== Detailed Accuracy By Class ====

          TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
  0.994     0.116     0.992     0.994     0.993     0.888     0.918     0.983   negative
  0.884     0.006     0.905     0.884     0.895     0.888     0.918     0.831   sick
Weighted Avg.  0.987     0.109     0.987     0.987     0.987     0.888     0.918     0.974

==== Confusion Matrix ====

      a      b  <-- classified as
3203  20 |      a = negative
      25  191 |      b = sick

```

## Percentage Split 50%

```

==== Run information ====

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3439
Attributes:  248
              [list of attributes omitted]
Test mode:   split 50.0% train, remainder test

==== Classifier model (full training set) ====

PART decision list
-----
T3_3 = '(1.15-inf)' AND
FTI_39 = '(-inf-170.5)' AND
TT4_1 = '(61.5-inf)': negative (3079.0/1.0)

T3_3 = '(1.15-inf)' AND
T4U_33 = '(-inf-1.115)' AND
referral_source = other AND
TSH_12 = '(0.765-inf)': negative (27.0)

T3_19 = '(-inf-2.65]' AND
T3_9 = '(-inf-1.75)' AND
TT4_39 = '(-inf-169.5)' AND
FTI_38 = '(-inf-157.5)' AND
TT4_37 = '(-inf-147.5)' AND
referral_source = SVI AND
TT4_1 = '(61.5-inf)': sick (146.0/1.0)

T3_19 = '(2.65-inf)': negative (17.0)

T4U_measured = t AND
age_37 = '(77.5-inf)' AND
TT4_32 = '(-inf-127.5)': sick (15.0)

T4U_measured = t AND
TT4_17 = '(101.5-inf)' AND
on_thyroxine = t: negative (24.0)

goitre = f AND
T4U_33 = '(-inf-1.115)' AND
T4U_measured = t AND
on_thyroxine = f AND
age_1 = '(17.5-inf)' AND
T3_measured = t AND
TSH_2 = '(0.0375-inf)' AND
sick = f AND
TSH_37 = '(-inf-6.75)' AND
T3_1 = '(0.75-inf)' AND
T4U_2 = '(0.745-inf)' AND
T4U_31 = '(-inf-1.075)' AND
T4U_27 = '(-inf-1.035)' AND
TSH_7 = '(0.295-inf)' AND
FTI_16 = '(104.5-inf)': negative (13.0)

T4U_measured = t AND
on_thyroxine = f AND
age_1 = '(17.5-inf)' AND
goitre = f AND
T3_17 = '(-inf-2.45]' AND
T4U_33 = '(-inf-1.115)' AND
T3_measured = t AND
TSH_2 = '(0.0375-inf)' AND
sick = f AND
age_3 = '(22.5-inf)' AND
T3_12 = '(-inf-2.00675)' AND
TSH_3 = '(0.075-inf)' AND
TSH_37 = '(-inf-6.75)' AND
age_34 = '(-inf-72.5)' AND
T3_9 = '(-inf-1.75)' AND
age_9 = '(34.5-inf)' AND
referral_source = other: negative (17.0/1.0)

T4U_measured = t AND
age_1 = '(17.5-inf)' AND
on_thyroxine = f AND
T4U_29 = '(-inf-1.055)' AND
T3_measured = f AND
sex = F: negative (9.0)

T4U_measured = t AND
age_1 = '(17.5-inf)' AND
TT4_38 = '(-inf-157.5)' AND

```

```

age_37 = '(-inf-77.5]' AND
T3_13 = '(-inf-2.05675]' AND
TSH_3 = '(0.075-inf)' AND
TT4_1 = '(61.5-inf)': sick (19.0/1.0)

TT4_38 = '(-inf-157.5]' AND
T4U_28 = '(-inf-1.045]' AND
goitre = f AND
T4U_13 = '(0.905-inf)': negative (14.0)

age_1 = '(17.5-inf)' AND
psych = f AND
T4U_11 = '(0.885-inf)': sick (13.0)

age_5 = '(26.5-inf)' AND
psych = f AND
TT4_38 = '(-inf-157.5]' AND
age_37 = '(-inf-77.5]' AND
T3_3 = '(1.15-inf)' AND
age_16 = '(-inf-48.5)': negative (5.0)

age_5 = '(26.5-inf)' AND
psych = f AND
age_15 = '(46.5-inf)' AND
TT4_38 = '(-inf-157.5]' AND
age_37 = '(-inf-77.5]' AND
TSH_14 = '(0.925-inf)' AND
TSH_28 = '(-inf-2.55)': sick (7.0)

TT4_38 = '(-inf-157.5]' AND
age_5 = '(26.5-inf)' AND
age_34 = '(-inf-72.5]' AND
psych = f AND
age_20 = '(-inf-55.5)': sick (10.0/1.0)

TT4_38 = '(-inf-157.5]' AND
age_28 = '(-inf-63.5)': negative (9.0)

TT4_38 = '(-inf-157.5]' AND
TSH_12 = '(-inf-0.765)': negative (9.0/1.0)

: sick (6.0)

Number of Rules : 18

Time taken to build model: 0.3 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0.01 seconds

==== Summary ===

    Correctly Classified Instances      1680           97.7312 %
    Incorrectly Classified Instances   39            2.2688 %
    Kappa statistic                   0.7958
    Mean absolute error              0.0239
    Root mean squared error          0.1468
    Relative absolute error          20.2384 %
    Root relative squared error     62.9546 %
    Total Number of Instances       1719

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.986     0.172     0.989     0.986     0.988     0.796     0.904     0.991     negative
      0.828     0.014     0.788     0.828     0.808     0.796     0.904     0.754     sick
Weighted Avg.   0.977     0.163     0.978     0.977     0.978     0.796     0.904     0.977

==== Confusion Matrix ===

      a     b  <-- classified as
1598   22 |   a = negative
      17   82 |   b = sick

```

## Percentage Split 60%

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3439
Attributes:  248

```

```

[list of attributes omitted]
Test mode: split 60.0% train, remainder test

==== Classifier model (full training set) ===

PART decision list
-----

T3_3 = '(1.15-inf)' AND
FTI_39 = '(-inf-170.5)' AND
TT4_1 = '(61.5-inf)': negative (3079.0/1.0)

T3_3 = '(1.15-inf)' AND
T4U_33 = '(-inf-1.115)' AND
referral_source = other AND
TSH_12 = '(0.765-inf)': negative (27.0)

T3_19 = '(-inf-2.65)' AND
T3_9 = '(-inf-1.75)' AND
TT4_39 = '(-inf-169.5)' AND
FTI_38 = '(-inf-157.5)' AND
TT4_37 = '(-inf-147.5)' AND
referral_source = SVI AND
TT4_1 = '(61.5-inf)': sick (146.0/1.0)

T3_19 = '(2.65-inf)': negative (17.0)

T4U_measured = t AND
age_37 = '(77.5-inf)' AND
TT4_32 = '(-inf-127.5)': sick (15.0)

T4U_measured = t AND
TT4_17 = '(101.5-inf)' AND
on_thyroxine = t: negative (24.0)

goitre = f AND
T4U_33 = '(-inf-1.115)' AND
T4U_measured = t AND
on_thyroxine = f AND
age_1 = '(17.5-inf)' AND
T3_measured = t AND
TSH_2 = '(0.0375-inf)' AND
sick = f AND
TSH_37 = '(-inf-6.75)' AND
T3_1 = '(0.75-inf)' AND
T4U_2 = '(0.745-inf)' AND
T4U_31 = '(-inf-1.075)' AND
T4U_27 = '(-inf-1.035)' AND
TSH_7 = '(0.295-inf)' AND
FTI_16 = '(104.5-inf)': negative (13.0)

T4U_measured = t AND
on_thyroxine = f AND
age_1 = '(17.5-inf)' AND
goitre = f AND
T3_17 = '(-inf-2.45)' AND
T4U_33 = '(-inf-1.115)' AND
T3_measured = t AND
TSH_2 = '(0.0375-inf)' AND
sick = f AND
age_3 = '(22.5-inf)' AND
T3_12 = '(-inf-2.00675)' AND
TSH_3 = '(0.075-inf)' AND
TSH_37 = '(-inf-6.75)' AND
age_34 = '(-inf-72.5)' AND
T3_9 = '(-inf-1.75)' AND
age_9 = '(34.5-inf)' AND
referral_source = other: negative (17.0/1.0)

T4U_measured = t AND
age_1 = '(17.5-inf)' AND
on_thyroxine = f AND
T4U_29 = '(-inf-1.055)' AND
T3_measured = f AND
sex = F: negative (9.0)

T4U_measured = t AND
age_1 = '(17.5-inf)' AND
TT4_38 = '(-inf-157.5)' AND
age_37 = '(-inf-77.5)' AND
T3_13 = '(-inf-2.05675)' AND
TSH_3 = '(0.075-inf)' AND
TT4_1 = '(61.5-inf)': sick (19.0/1.0)

TT4_38 = '(-inf-157.5)' AND
T4U_28 = '(-inf-1.045)' AND
goitre = f AND

```

```

T4U_13 = '(0.905-inf)': negative (14.0)
age_1 = '(17.5-inf)' AND
psych = f AND
T4U_11 = '(0.885-inf)': sick (13.0)

age_5 = '(26.5-inf)' AND
psych = f AND
TT4_38 = '(-inf-157.5)' AND
age_37 = '(-inf-77.5]' AND
T3_3 = '(1.15-inf)' AND
age_16 = '(-inf-48.5]': negative (5.0)

age_5 = '(26.5-inf)' AND
psych = f AND
age_15 = '(46.5-inf)' AND
TT4_38 = '(-inf-157.5)' AND
age_37 = '(-inf-77.5]' AND
TSH_14 = '(0.925-inf)' AND
TSH_28 = '(-inf-2.55]': sick (7.0)

TT4_38 = '(-inf-157.5)' AND
age_5 = '(26.5-inf)' AND
age_34 = '(-inf-72.5]' AND
psych = f AND
age_20 = '(-inf-55.5]': sick (10.0/1.0)

TT4_38 = '(-inf-157.5)' AND
age_28 = '(-inf-63.5]': negative (9.0)

TT4_38 = '(-inf-157.5)' AND
TSH_12 = '(-inf-0.765)': negative (9.0/1.0)

: sick (6.0)

Number of Rules : 18

Time taken to build model: 0.31 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0.01 seconds

==== Summary ===

    Correctly Classified Instances      1343          97.6017 %
    Incorrectly Classified Instances     33           2.3983 %
    Kappa statistic                      0.792
    Mean absolute error                  0.0249
    Root mean squared error              0.1501
    Relative absolute error              20.9206 %
    Root relative squared error         60.6922 %
    Total Number of Instances            1376

==== Detailed Accuracy By Class ===

          TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
          0.991     0.244     0.983     0.991     0.987     0.794     0.928     0.992     negative
          0.756     0.009     0.861     0.756     0.805     0.794     0.928     0.797     sick
Weighted Avg.   0.976     0.229     0.975     0.976     0.975     0.794     0.928     0.979

==== Confusion Matrix ===

      a      b  <-- classified as
1275   11 |   a = negative
      22   68 |   b = sick

```

## Percentage Split 66%

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3439
Attributes:  248
              [list of attributes omitted]
Test mode:   split 66.0% train, remainder test

==== Classifier model (full training set) ===

PART decision list
-----

```

```

T3_3 = '(1.15-inf)' AND
FTI_39 = '(-inf-170.5]' AND
TT4_1 = '(61.5-inf)': negative (3079.0/1.0)

T3_3 = '(1.15-inf)' AND
T4U_33 = '(-inf-1.115]' AND
referral_source = other AND
TSH_12 = '(0.765-inf)': negative (27.0)

T3_19 = '(-inf-2.65]' AND
T3_9 = '(-inf-1.75]' AND
TT4_39 = '(-inf-169.5]' AND
FTI_38 = '(-inf-157.5]' AND
TT4_37 = '(-inf-147.5]' AND
referral_source = SVI AND
TT4_1 = '(61.5-inf)': sick (146.0/1.0)

T3_19 = '(2.65-inf)': negative (17.0)

T4U_measured = t AND
age_37 = '(77.5-inf)' AND
TT4_32 = '(-inf-127.5)': sick (15.0)

T4U_measured = t AND
TT4_17 = '(101.5-inf)' AND
on_thyroxine = t: negative (24.0)

goitre = f AND
T4U_33 = '(-inf-1.115]' AND
T4U_measured = t AND
on_thyroxine = f AND
age_1 = '(17.5-inf)' AND
T3_measured = t AND
TSH_2 = '(0.0375-inf)' AND
sick = f AND
TSH_37 = '(-inf-6.75]' AND
T3_1 = '(0.75-inf)' AND
T4U_2 = '(0.745-inf)' AND
T4U_31 = '(-inf-1.075]' AND
T4U_27 = '(-inf-1.035]' AND
TSH_7 = '(0.295-inf)' AND
FTI_16 = '(104.5-inf)': negative (13.0)

T4U_measured = t AND
on_thyroxine = f AND
age_1 = '(17.5-inf)' AND
goitre = f AND
T3_17 = '(-inf-2.45]' AND
T4U_33 = '(-inf-1.115]' AND
T3_measured = t AND
TSH_2 = '(0.0375-inf)' AND
sick = f AND
age_3 = '(22.5-inf)' AND
T3_12 = '(-inf-2.00675]' AND
TSH_3 = '(0.075-inf)' AND
TSH_37 = '(-inf-6.75]' AND
age_34 = '(-inf-72.5]' AND
T3_9 = '(-inf-1.75]' AND
age_9 = '(34.5-inf)' AND
referral_source = other: negative (17.0/1.0)

T4U_measured = t AND
age_1 = '(17.5-inf)' AND
on_thyroxine = f AND
T4U_29 = '(-inf-1.055]' AND
T3_measured = f AND
sex = F: negative (9.0)

T4U_measured = t AND
age_1 = '(17.5-inf)' AND
TT4_38 = '(-inf-157.5]' AND
age_37 = '(-inf-77.5]' AND
T3_13 = '(-inf-2.05675]' AND
TSH_3 = '(0.075-inf)' AND
TT4_1 = '(61.5-inf)': sick (19.0/1.0)

TT4_38 = '(-inf-157.5]' AND
T4U_28 = '(-inf-1.045]' AND
goitre = f AND
T4U_13 = '(0.905-inf)': negative (14.0)

age_1 = '(17.5-inf)' AND
psych = f AND
T4U_11 = '(0.885-inf)': sick (13.0)

age_5 = '(26.5-inf)' AND
psych = f AND

```

```

TT4_38 = '(-inf-157.5)' AND
age_37 = '(-inf-77.5)' AND
T3_3 = '(1.15-inf)' AND
age_16 = '(-inf-48.5)': negative (5.0)

age_5 = '(26.5-inf)' AND
psych = f AND
age_15 = '(46.5-inf)' AND
TT4_38 = '(-inf-157.5)' AND
age_37 = '(-inf-77.5)' AND
TSH_14 = '(0.925-inf)' AND
TSH_28 = '(-inf-2.55)': sick (7.0)

TT4_38 = '(-inf-157.5)' AND
age_5 = '(26.5-inf)' AND
age_34 = '(-inf-72.5)' AND
psych = f AND
age_20 = '(-inf-55.5)': sick (10.0/1.0)

TT4_38 = '(-inf-157.5)' AND
age_28 = '(-inf-63.5)': negative (9.0)

TT4_38 = '(-inf-157.5)' AND
TSH_12 = '(-inf-0.765)': negative (9.0/1.0)

: sick (6.0)

Number of Rules : 18

Time taken to build model: 0.32 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0.02 seconds

==== Summary ===

Correctly Classified Instances      1137           97.2626 %
Incorrectly Classified Instances   32            2.7374 %
Kappa statistic                   0.7536
Mean absolute error               0.0274
Root mean squared error          0.1585
Relative absolute error          23.5056 %
Root relative squared error     67.2367 %
Total Number of Instances        1169

==== Detailed Accuracy By Class ===

          TP Rate  FP Rate  Precision  Recall  F-Measure  MCC    ROC Area  PRC Area  Class
          0.985    0.232    0.985    0.985    0.985    0.754    0.866    0.985    negative
          0.768    0.015    0.768    0.768    0.768    0.754    0.866    0.734    sick
Weighted Avg.    0.973    0.219    0.973    0.973    0.973    0.754    0.866    0.970

==== Confusion Matrix ===

      a      b  <-- classified as
1084  16 |      a = negative
      16  53 |      b = sick

```

## Leave one out fold

```

==== Run information ===

Scheme:      weka.classifiers.rules.PART -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3439
Attributes:  248
              [list of attributes omitted]
Test mode:   3439-fold cross-validation

==== Classifier model (full training set) ===

PART decision list
-----
T3_3 = '(1.15-inf)' AND
FTI_39 = '(-inf-170.5)' AND
TT4_1 = '(61.5-inf)': negative (3079.0/1.0)

T3_3 = '(1.15-inf)' AND
T4U_33 = '(-inf-1.115)' AND
referral_source = other AND
TSH_12 = '(0.765-inf)': negative (27.0)

```

```

T3_19 = '(-inf-2.65]' AND
T3_9 = '(-inf-1.75]' AND
TT4_39 = '(-inf-169.5]' AND
FT4_38 = '(-inf-157.5]' AND
TT4_37 = '(-inf-147.5]' AND
referral_source = SVI AND
TT4_1 = '(61.5-inf)': sick (146.0/1.0)

T3_19 = '(2.65-inf)': negative (17.0)

T4U_measured = t AND
age_37 = '(77.5-inf)' AND
TT4_32 = '(-inf-127.5)': sick (15.0)

T4U_measured = t AND
TT4_17 = '(101.5-inf)' AND
on_thyroxine = t: negative (24.0)

goitre = f AND
T4U_33 = '(-inf-1.115]' AND
T4U_measured = t AND
on_thyroxine = f AND
age_1 = '(17.5-inf)' AND
T3_measured = t AND
TSH_2 = '(0.0375-inf)' AND
sick = f AND
TSH_37 = '(-inf-6.75]' AND
T3_1 = '(0.75-inf)' AND
T4U_2 = '(0.745-inf)' AND
T4U_31 = '(-inf-1.075]' AND
T4U_27 = '(-inf-1.035)' AND
TSH_7 = '(0.295-inf)' AND
FTI_16 = '(104.5-inf)': negative (13.0)

T4U_measured = t AND
on_thyroxine = f AND
age_1 = '(17.5-inf)' AND
goitre = f AND
T3_17 = '(-inf-2.45]' AND
T4U_33 = '(-inf-1.115]' AND
T3_measured = t AND
TSH_2 = '(0.0375-inf)' AND
sick = f AND
age_3 = '(22.5-inf)' AND
T3_12 = '(-inf-2.00675]' AND
TSH_3 = '(0.075-inf)' AND
TSH_37 = '(-inf-6.75]' AND
age_34 = '(-inf-72.5]' AND
T3_9 = '(-inf-1.75]' AND
age_9 = '(34.5-inf)' AND
referral_source = other: negative (17.0/1.0)

T4U_measured = t AND
age_1 = '(17.5-inf)' AND
on_thyroxine = f AND
T4U_29 = '(-inf-1.055)' AND
T3_measured = f AND
sex = F: negative (9.0)

T4U_measured = t AND
age_1 = '(17.5-inf)' AND
TT4_38 = '(-inf-157.5]' AND
age_37 = '(-inf-77.5]' AND
T3_13 = '(-inf-2.05675]' AND
TSH_3 = '(0.075-inf)' AND
TT4_1 = '(61.5-inf)': sick (19.0/1.0)

TT4_38 = '(-inf-157.5]' AND
T4U_28 = '(-inf-1.045)' AND
goitre = f AND
T4U_13 = '(0.905-inf)': negative (14.0)

age_1 = '(17.5-inf)' AND
psych = f AND
T4U_11 = '(0.885-inf)': sick (13.0)

age_5 = '(26.5-inf)' AND
psych = f AND
TT4_38 = '(-inf-157.5)' AND
age_37 = '(-inf-77.5)' AND
T3_3 = '(1.15-inf)' AND
age_16 = '(-inf-48.5)': negative (5.0)

age_5 = '(26.5-inf)' AND
psych = f AND
age_15 = '(46.5-inf)' AND

```

```

TT4_38 = '(-inf-157.5)' AND
age_37 = '(-inf-77.5)' AND
TSH_14 = '(0.925-inf)' AND
TSH_28 = '(-inf-2.55)': sick (7.0)

TT4_38 = '(-inf-157.5)' AND
age_5 = '(26.5-inf)' AND
age_34 = '(-inf-72.5)' AND
psych = f AND
age_20 = '(-inf-55.5)': sick (10.0/1.0)

TT4_38 = '(-inf-157.5)' AND
age_28 = '(-inf-63.5)': negative (9.0)

TT4_38 = '(-inf-157.5)' AND
TSH_12 = '(-inf-0.765)': negative (9.0/1.0)

: sick (6.0)

Number of Rules : 18

Time taken to build model: 0.32 seconds

==== Stratified cross-validation ====
==== Summary ====

    Correctly Classified Instances      3386          98.4589 %
    Incorrectly Classified Instances   53           1.5411 %
    Kappa statistic                   0.8682
    Mean absolute error              0.0165
    Root mean squared error         0.1213
    Relative absolute error         14.0193 %
    Root relative squared error     49.983 %
    Total Number of Instances       3439

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.992     0.130     0.991     0.992     0.992     0.868     0.891     0.987     negative
      0.870     0.008     0.883     0.870     0.876     0.868     0.891     0.790     sick
Weighted Avg.   0.985     0.122     0.984     0.985     0.985     0.868     0.891     0.975

==== Confusion Matrix ====

      a     b  <-- classified as
3198   25 |     a = negative
      28  188 |     b = sick

```

### 3.6.3 Conclusion

Preprocessing method	PART							
	Cross Validation			Percentage Split			Leave One Out Fold	
	5 folds	10 folds	15 folds	50.00%	60.00%	66.00%		
With Original Dataset	98.3033%	98.6214%	98.5949%	98.6214%	97.8131%	98.7520%	98.8865%	
Missing + Extreme + Outliers	98.6624%	98.4879%	98.8950%	98.6038%	98.3285%	97.6903%	98.8078%	
Missing + Extreme + Outliers + frequency 40	98.2262%	98.4879%	98.6915%	97.7312%	97.6017%	97.2626%	98.4589%	

Dealing with missing, extreme and outlier values was a good combo as it was able to win in 3 tests, in Cross validation 5 and 15 folds, also in Percentage split 60%. Adding to the combo discretizing with 40 bins which also showed good results yielded the opposed outcome and decreased the performance not winning in any test.

## 3.7 General Conclusion

		Preprocessing method		PART			Leave One Out Fold	AVG	Order AVG	TOP 3 results
				Cross Validation	Percentage Split					
Discretization	Non supervised	With Original Dataset		5 folds	10 folds	15 folds	50.00%	60.00%	66.00%	98.8865% 98.5132% 1 3
		Equal width binning		40 bins	98.3298%	98.4889%	98.4093%	98.6214%	97.8131%	98.7520%
		5 bins		10 bins	98.2238%	97.9321%	98.1177%	97.4549%	98.0782%	97.8939%
		2 bins		5 bins	93.8494%	94.1676%	94.1145%	93.7964%	94.0358%	94.4618%
		Equal frequency binning		40 bins	93.7434%	93.8494%	93.8229%	93.7434%	94.1021%	93.2917%
		5 bins		10 bins	98.5419%	98.2768%	97.9586%	98.6214%	98.0782%	98.2839%
		2 bins		5 bins	98.1442%	97.7466%	97.9056%	97.8261%	97.2830%	97.9719%
		Supervised		2 bins	93.7699%	93.4517%	93.7434%	93.6373%	93.9032%	94.0718%
		With Filter		Supervised	98.0382%	97.9586%	97.9586%	97.6140%	97.9457%	97.7379%
		With Wrapper		With Filter	98.1972%	98.2238%	98.0117%	98.0912%	98.1445%	97.7379%
Feature Selection	Missing Values	Missing Values only		With Wrapper	98.1442%	98.1177%	98.1707%	97.4549%	97.3492%	98.4399%
		Missing Values + Extreme Values		Missing Values only	98.1177%	98.7805%	98.7010%	96.9777%	97.8131%	98.9860%
		Missing Values + Outliers		Missing Values + Extreme Values	98.2079%	98.3457%	98.3182%	98.2350%	98.3460%	98.2157%
		Missing + Extreme + Outliers		Missing Values + Outliers	98.3089%	98.5626%	98.5062%	98.3653%	98.1672%	98.3416%
Combos		Missing + Extreme + Outliers		Missing + Extreme + Outliers	98.6624%	98.4879%	98.8950%	98.6038%	98.3285%	97.6903%
		Missing + Extreme + Outliers + frequency 40		Missing + Extreme + Outliers + frequency 40	98.2262%	98.4879%	98.6915%	97.7312%	97.6017%	97.2626%

Taking a look at the big picture we see that just running PART with the original dataset gives great performance and it is very hard to find measure to the dataset to gain improvements in performance. if we look at the best 3 results we find that we find that dealing with missing values with the test of Percentage split 66% we get **98.9860%**. after that the second best result is given in dealing with missing , extreme and outlier values with the test of Cross Validation and 15 folds with **98.8950%**. third place is for just running the original dataset with leave one out fold we get **98.8865%** correctly classified instances.

However if we take the average of each modification across the entire tests we find that just running PART with the original dataset **98.5132%**. and in second place we see that dealing with missing , extreme and outlier values gives an average of **98.4965%**

## Chapter 4

# Testing with preprocessed dataset and C4.5 Algorithm

## 4.1 Non supervised discretization and Equal width binning

### 4.1.1 2 bins

#### Cross Validation 5 folds

```
==== Run information ====

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B2-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  30
              age_1
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              T3_measured
              T3_1
              TT4_measured
              TT4_1
              T4U_measured
              T4U_1
              FTI_measured
              FTI_1
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   5-fold cross-validation

==== Classifier model (full training set) ====

J48 pruned tree
-----
: negative (3772.0/231.0)

Number of Leaves : 1

Size of the tree : 1

Time taken to build model: 0.11 seconds

==== Stratified cross-validation ====
==== Summary ===
```

```

Correctly Classified Instances      3541          93.8759 %
Incorrectly Classified Instances   231           6.1241 %
Kappa statistic                   0
Mean absolute error               0.115
Root mean squared error          0.2398
Relative absolute error           99.7787 %
Root relative squared error      99.9999 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC     ROC Area   PRC Area   Class
      1.000     1.000     0.939      1.000     0.968     ?       0.498     0.939    negative
      0.000     0.000     ?          0.000     ?         ?       0.498     0.061    sick
Weighted Avg.      0.939     0.939     ?          0.939     ?         ?       0.498     0.885

==== Confusion Matrix ====

      a      b  <-- classified as
3541  0 |  a = negative
 231  0 |  b = sick

```

## Cross Validation 10 folds

```

==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B2-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  30
              age_1
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              T3_measured
              T3_1
              TT4_measured
              TT4_1
              T4U_measured
              T4U_1
              FTI_measured
              FTI_1
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   10-fold cross-validation

==== Classifier model (full training set) ===

J48 pruned tree
-----
: negative (3772.0/231.0)

Number of Leaves : 1

Size of the tree : 1

Time taken to build model: 0.09 seconds

==== Stratified cross-validation ===
==== Summary ===

Correctly Classified Instances      3541          93.8759 %
Incorrectly Classified Instances   231           6.1241 %
Kappa statistic                   0
Mean absolute error               0.115
Root mean squared error          0.2398
Relative absolute error           99.8032 %

```

```

Root relative squared error          99.9999 %
Total Number of Instances          3772

==== Detailed Accuracy By Class ====
      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
      1.000    1.000    0.939     1.000    0.968     ?       0.498    0.938    negative
      0.000    0.000    ?         0.000    ?         ?       0.498    0.061    sick
Weighted Avg.    0.939    0.939    ?         0.939    ?         ?       0.498    0.885

==== Confusion Matrix ====
      a      b  <-- classified as
3541  0 |  a = negative
  231  0 |  b = sick

```

## Cross Validation 15 folds

```

==== Run information ====
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B2-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  30
              age_1
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              t131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              T3_measured
              T3_1
              TT4_measured
              TT4_1
              T4U_measured
              T4U_1
              FTI_measured
              FTI_1
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   15-fold cross-validation

==== Classifier model (full training set) ====
J48 pruned tree
-----
: negative (3772.0/231.0)

Number of Leaves : 1
Size of the tree : 1

Time taken to build model: 0.08 seconds

==== Stratified cross-validation ====
==== Summary ===

  Correctly Classified Instances      3541           93.8759 %
  Incorrectly Classified Instances    231            6.1241 %
  Kappa statistic                   0
  Mean absolute error               0.115
  Root mean squared error          0.2398
  Relative absolute error          99.8102 %
  Root relative squared error      99.9999 %
  Total Number of Instances        3772

==== Detailed Accuracy By Class ====
      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class

```

```

    1.000    1.000    0.939    1.000    0.968    ?    0.492    0.938    negative
    0.000    0.000    ?        0.000    ?        ?    0.492    0.060    sick
Weighted Avg.    0.939    0.939    ?        0.939    ?        ?    0.492    0.884

==== Confusion Matrix ====
a      b    <-- classified as
3541   0 |     a = negative
231    0 |     b = sick

```

## Percentage Split 50%

```

==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B2-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  30
              age_1
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              i131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              T3_measured
              T3_1
              TT4_measured
              TT4_1
              T4U_measured
              T4U_1
              FTI_measured
              FTI_1
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 50.0% train, remainder test

==== Classifier model (full training set) ===

J48 pruned tree
-----
: negative (3772.0/231.0)

Number of Leaves : 1

Size of the tree : 1

Time taken to build model: 0.08 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

==== Summary ===

Correctly Classified Instances      1769           93.7964 %
Incorrectly Classified Instances   117            6.2036 %
Kappa statistic                   0
Mean absolute error               0.115
Root mean squared error          0.2412
Relative absolute error           99.6465 %
Root relative squared error     100.0011 %
Total Number of Instances        1886

==== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
      1.000    1.000    0.938     1.000    0.968     ?      0.500    0.938    negative
      0.000    0.000    ?        0.000    ?        ?      0.500    0.062    sick
Weighted Avg.    0.938    0.938    ?        0.938    ?        ?      0.500    0.884

```

```
==== Confusion Matrix ====
      a     b  <-- classified as
1769    0 |     a = negative
   17    0 |     b = sick
```

## Percentage Split 60%

```
==== Run information ====
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B2-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  30
              age_1
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              T3_measured
              T3_1
              TT4_measured
              TT4_1
              T4U_measured
              T4U_1
              FTI_measured
              FTI_1
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 60.0% train, remainder test

==== Classifier model (full training set) ====
J48 pruned tree
-----
: negative (3772.0/231.0)

Number of Leaves : 1
Size of the tree : 1

Time taken to build model: 0.09 seconds

==== Evaluation on test split ====
Time taken to test model on test split: 0.01 seconds

==== Summary ====
Correctly Classified Instances      1420           94.1021 %
Incorrectly Classified Instances    89            5.8979 %
Kappa statistic                   0
Mean absolute error               0.1143
Root mean squared error          0.2356
Relative absolute error           99.703 %
Root relative squared error      99.9972 %
Total Number of Instances        1509

==== Detailed Accuracy By Class ====
      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
      1.000    1.000    0.941    1.000    0.970    ?       0.500    0.941    negative
      0.000    0.000    ?        0.000    ?        ?       0.500    0.059    sick
Weighted Avg.  0.941    0.941    ?        0.941    ?        ?       0.500    0.889

==== Confusion Matrix ====

```

```
a      b    <-- classified as
1420   0 |    a = negative
  89   0 |    b = sick
```

## Percentage Split 66%

```
==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B2-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  30
              age_1
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              i131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              T3_measured
              T3_1
              TT4_measured
              TT4_1
              T4U_measured
              T4U_1
              FTI_measured
              FTI_1
              TBG_measured
              TBG_
              referral_source
              Class
Test mode:   split 66.0% train, remainder test

==== Classifier model (full training set) ===

J48 pruned tree
-----
: negative (3772.0/231.0)

Number of Leaves : 1
Size of the tree : 1

Time taken to build model: 0.1 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0.01 seconds

==== Summary ===

  Correctly Classified Instances      1207           94.1498 %
  Incorrectly Classified Instances    75            5.8502 %
  Kappa statistic                   0
  Mean absolute error               0.1138
  Root mean squared error          0.2347
  Relative absolute error          99.7284 %
  Root relative squared error     99.9972 %
  Total Number of Instances        1282

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      1.000     1.000     0.941      1.000     0.970      ?       0.500     0.941     negative
      0.000     0.000     ?          0.000     ?          ?       0.500     0.059     sick
  Weighted Avg.   0.941     0.941     ?          0.941     ?          ?       0.500     0.890

==== Confusion Matrix ===

a      b    <-- classified as
1207   0 |    a = negative
  75   0 |    b = sick
```

## Leave one out fold

```

==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B2-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  30
              age_1
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              T3_measured
              T3_1
              TT4_measured
              TT4_1
              T4U_measured
              T4U_1
              FTI_measured
              FTI_1
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   3772-fold cross-validation

==== Classifier model (full training set) ===

J48 pruned tree
-----
: negative (3772.0/231.0)

Number of Leaves : 1
Size of the tree : 1

Time taken to build model: 0.09 seconds

==== Stratified cross-validation ===
==== Summary ===

  Correctly Classified Instances      3541           93.8759 %
  Incorrectly Classified Instances    231            6.1241 %
  Kappa statistic                   0
  Mean absolute error               0.115
  Root mean squared error          0.2398
  Relative absolute error          99.8229 %
  Root relative squared error     100            %
  Total Number of Instances        3772

==== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
      1.000    1.000    0.939     1.000   0.968     ?       0.000    0.939    negative
      0.000    0.000    ?         0.000   ?         ?       0.000    0.061    sick
Weighted Avg.  0.939    0.939    ?         0.939   ?         ?       0.000    0.885

==== Confusion Matrix ===

  a      b  <-- classified as
3541    0 |  a = negative
  231    0 |  b = sick

```

## 4.1.2 5 bins

### Cross Validation 5 folds

```

==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B5-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  48
              age_1
              age_2
              age_3
              age_4
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              TSH_2
              TSH_3
              TSH_4
              T3_measured
              T3_1
              T3_2
              T3_3
              T3_4
              TT4_measured
              TT4_1
              TT4_2
              TT4_3
              TT4_4
              T4U_measured
              T4U_1
              T4U_2
              T4U_3
              T4U_4
              FTI_measured
              FTI_1
              FTI_2
              FTI_3
              FTI_4
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   5-fold cross-validation

==== Classifier model (full training set) ===

J48 pruned tree
-----
: negative (3772.0/231.0)

Number of Leaves : 1
Size of the tree : 1

Time taken to build model: 0.18 seconds

==== Stratified cross-validation ===
==== Summary ===

  Correctly Classified Instances      3541          93.8759 %
  Incorrectly Classified Instances    231           6.1241 %
  Kappa statistic                   0
  Mean absolute error               0.115
  Root mean squared error          0.2398
  Relative absolute error          99.7787 %
  Root relative squared error     99.9999 %
  Total Number of Instances        3772

==== Detailed Accuracy By Class ===

```

```

TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
1.000    1.000    0.939     1.000    0.968     ?       0.498     0.939    negative
0.000    0.000    ?         0.000    ?         ?       0.498     0.061    sick
Weighted Avg. 0.939    0.939    ?         0.939    ?         ?       0.498     0.885

==== Confusion Matrix ====

a      b  <-- classified as
3541   0 |   a = negative
231    0 |   b = sick

```

## Cross Validation 10 folds

```

==== Run information ===

Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2
Relation: sick-weka.filters.unsupervised.attribute.Discretize-D-B5-M-1.0-Rfirst-last-precision6
Instances: 3772
Attributes: 48
age_1
age_2
age_3
age_4
sex
on_thyroxine
query_on_thyroxine
on_antithyroid_medication
sick
pregnant
thyroid_surgery
I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH_1
TSH_2
TSH_3
TSH_4
T3_measured
T3_1
T3_2
T3_3
T3_4
TT4_measured
TT4_1
TT4_2
TT4_3
TT4_4
T4U_measured
T4U_1
T4U_2
T4U_3
T4U_4
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
TBG_measured
TBG
referral_source
Class
Test mode: 10-fold cross-validation

==== Classifier model (full training set) ===

J48 pruned tree
-----
: negative (3772.0/231.0)

Number of Leaves : 1

Size of the tree : 1

Time taken to build model: 0.14 seconds

==== Stratified cross-validation ===

```

```

==== Summary ====
Correctly Classified Instances      3541          93.8759 %
Incorrectly Classified Instances   231           6.1241 %
Kappa statistic                   0
Mean absolute error               0.115
Root mean squared error          0.2398
Relative absolute error           99.8032 %
Root relative squared error     99.9999 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC    ROC Area  PRC Area  Class
1.000     1.000     0.939     1.000     0.968     ?       0.498     0.938  negative
0.000     0.000     ?         0.000     ?         ?       0.498     0.061  sick
Weighted Avg.  0.939     0.939     ?         0.939     ?         ?       0.498     0.885

==== Confusion Matrix ====

      a      b  <-- classified as
3541  0 |  a = negative
231   0 |  b = sick

```

## Cross Validation 15 folds

```

==== Run information ====

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B5-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  48
              age_1
              age_2
              age_3
              age_4
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              TSH_2
              TSH_3
              TSH_4
              T3_measured
              T3_1
              T3_2
              T3_3
              T3_4
              TT4_measured
              TT4_1
              TT4_2
              TT4_3
              TT4_4
              T4U_measured
              T4U_1
              T4U_2
              T4U_3
              T4U_4
              FTI_measured
              FTI_1
              FTI_2
              FTI_3
              FTI_4
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   15-fold cross-validation

==== Classifier model (full training set) ====

```

```
J48 pruned tree
-----
: negative (3772.0/231.0)

Number of Leaves : 1

Size of the tree : 1

Time taken to build model: 0.12 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3541          93.8759 %
Incorrectly Classified Instances   231           6.1241 %
Kappa statistic                      0
Mean absolute error                  0.115
Root mean squared error              0.2398
Relative absolute error              99.8102 %
Root relative squared error        99.9999 %
Total Number of Instances          3772

==== Detailed Accuracy By Class ====

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC    ROC Area  PRC Area  Class
1.000     1.000     0.939     1.000     0.968     ?       0.492     0.938  negative
0.000     0.000     ?         0.000     ?         ?       0.492     0.060  sick
Weighted Avg.   0.939     0.939     ?         0.939     ?         ?       0.492     0.884

==== Confusion Matrix ====

      a      b  <-- classified as
3541  0 |  a = negative
231   0 |  b = sick
```

## Percentage Split 50%

```
==== Run information ====

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B5-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  48
             age_1
             age_2
             age_3
             age_4
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             I131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary
             psych
             TSH_measured
             TSH_1
             TSH_2
             TSH_3
             TSH_4
             T3_measured
             T3_1
             T3_2
             T3_3
             T3_4
             TT4_measured
             TT4_1
             TT4_2
             TT4_3
             TT4_4
             T4U_measured
             T4U_1
             T4U_2
             T4U_3
             T4U_4
             FTI_measured
```

```

FTI_1
FTI_2
FTI_3
FTI_4
TBG_measured
TBG
referral_source
Class
Test mode: split 50.0% train, remainder test

==== Classifier model (full training set) ===

J48 pruned tree
-----
: negative (3772.0/231.0)

Number of Leaves : 1

Size of the tree : 1

Time taken to build model: 0.15 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

==== Summary ===

Correctly Classified Instances      1769          93.7964 %
Incorrectly Classified Instances   117           6.2036 %
Kappa statistic                      0
Mean absolute error                 0.115
Root mean squared error             0.2412
Relative absolute error              99.6465 %
Root relative squared error        100.0011 %
Total Number of Instances          1886

==== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC    ROC Area  PRC Area  Class
      1.000     1.000     0.938     1.000     0.968     ?      0.500     0.938  negative
      0.000     0.000     ?         0.000     ?         ?      0.500     0.062    sick
Weighted Avg.     0.938     0.938     ?         0.938     ?         ?      0.500     0.884

==== Confusion Matrix ===

      a      b  <-- classified as
1769  0 |  a = negative
  117  0 |  b = sick

```

## Percentage Split 60%

```

==== Run information ===

Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2
Relation: sick-weka.filters.unsupervised.attribute.Discretize-D-B5-M-1.0-Rfirst-last-precision6
Instances: 3772
Attributes: 48
age_1
age_2
age_3
age_4
sex
on_thyroxine
query_on_thyroxine
on_antithyroid_medication
sick
pregnant
thyroid_surgery
I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH_1
TSH_2
TSH_3
TSH_4
T3_measured

```

```

T3_1
T3_2
T3_3
T3_4
TT4_measured
TT4_1
TT4_2
TT4_3
TT4_4
T4U_measured
T4U_1
T4U_2
T4U_3
T4U_4
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
TBG_measured
TBG
referral_source
Class
Test mode: split 60.0% train, remainder test

==== Classifier model (full training set) ===

J48 pruned tree
-----
: negative (3772.0/231.0)

Number of Leaves : 1

Size of the tree : 1

Time taken to build model: 0.15 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

==== Summary ===

    Correctly Classified Instances      1420           94.1021 %
    Incorrectly Classified Instances     89            5.8979 %
    Kappa statistic                      0
    Mean absolute error                  0.1143
    Root mean squared error              0.2356
    Relative absolute error              99.703 %
    Root relative squared error          99.9972 %
    Total Number of Instances           1509

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC   ROC Area   PRC Area   Class
      1.000     1.000     0.941      1.000     0.970      ?     0.500      0.941   negative
      0.000     0.000     ?          0.000     ?          ?     0.500      0.059   sick
Weighted Avg.   0.941     0.941     ?          0.941     ?          ?     0.500      0.889

==== Confusion Matrix ===

  a     b  <-- classified as
1420   0 |   a = negative
  89   0 |   b = sick

```

## Percentage Split 66%

```

==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B5-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  48
             age_1
             age_2
             age_3
             age_4
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant

```

```

thyroid_surgery
I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH_1
TSH_2
TSH_3
TSH_4
T3_measured
T3_1
T3_2
T3_3
T3_4
TT4_measured
TT4_1
TT4_2
TT4_3
TT4_4
T4U_measured
T4U_1
T4U_2
T4U_3
T4U_4
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
TBG_measured
TBG
referral_source
Class
Test mode: split 66.0% train, remainder test

==== Classifier model (full training set) ====
J48 pruned tree
-----
: negative (3772.0/231.0)

Number of Leaves : 1
Size of the tree : 1

Time taken to build model: 0.15 seconds

==== Evaluation on test split ====
Time taken to test model on test split: 0.01 seconds

==== Summary ====
Correctly Classified Instances      1207      94.1498 %
Incorrectly Classified Instances    75       5.8502 %
Kappa statistic                      0
Mean absolute error                  0.1138
Root mean squared error              0.2347
Relative absolute error              99.7284 %
Root relative squared error         99.9972 %
Total Number of Instances           1282

==== Detailed Accuracy By Class ====
      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC   ROC Area   PRC Area   Class
      1.000     1.000     0.941      1.000     0.970      ?     0.500     0.941   negative
      0.000     0.000     ?          0.000     ?          ?     0.500     0.059   sick
Weighted Avg.     0.941     0.941     ?          0.941     ?          ?     0.500     0.890

==== Confusion Matrix ====
      a     b  <-- classified as
1207   0 |     a = negative
      75   0 |     b = sick

```

## Leave one out fold

```
==== Run information ====
```

```

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B5-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  48
             age_1
             age_2
             age_3
             age_4
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             I131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary
             psych
             TSH_measured
             TSH_1
             TSH_2
             TSH_3
             TSH_4
             T3_measured
             T3_1
             T3_2
             T3_3
             T3_4
             TT4_measured
             TT4_1
             TT4_2
             TT4_3
             TT4_4
             T4U_measured
             T4U_1
             T4U_2
             T4U_3
             T4U_4
             FTI_measured
             FTI_1
             FTI_2
             FTI_3
             FTI_4
             TBG_measured
             TBG
             referral_source
             Class
Test mode:   3772-fold cross-validation

==== Classifier model (full training set) ====
J48 pruned tree
-----
: negative (3772.0/231.0)

Number of Leaves : 1

Size of the tree : 1

Time taken to build model: 0.1 seconds

==== Stratified cross-validation ====
==== Summary ===

  Correctly Classified Instances      3541          93.8759 %
  Incorrectly Classified Instances   231           6.1241 %
  Kappa statistic                   0
  Mean absolute error              0.115
  Root mean squared error          0.2398
  Relative absolute error          99.8229 %
  Root relative squared error     100           %
  Total Number of Instances        3772

==== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
      1.000    1.000    0.939     1.000    0.968     ?       0.000    0.939    negative
      0.000    0.000    ?         0.000    ?         ?       0.000    0.061    sick
Weighted Avg.  0.939    0.939    ?         0.939    ?         ?       0.000    0.885

```

```
==== Confusion Matrix ====
      a     b  <-- classified as
3541   0 |    a = negative
 231   0 |    b = sick
```

### 4.1.3 10 bins

#### Cross Validation 5 folds

```
==== Run information ====
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B10-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  78
              age_1
              age_2
              age_3
              age_4
              age_5
              age_6
              age_7
              age_8
              age_9
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              TSH_2
              TSH_3
              TSH_4
              TSH_5
              TSH_6
              TSH_7
              TSH_8
              TSH_9
              T3_measured
              T3_1
              T3_2
              T3_3
              T3_4
              T3_5
              T3_6
              T3_7
              T3_8
              T3_9
              TT4_measured
              TT4_1
              TT4_2
              TT4_3
              TT4_4
              TT4_5
              TT4_6
              TT4_7
              TT4_8
              TT4_9
              T4U_measured
              T4U_1
              T4U_2
              T4U_3
              T4U_4
              T4U_5
              T4U_6
              T4U_7
              T4U_8
              T4U_9
              FTI_measured
              FTI_1
              FTI_2
```

```

FTI_3
FTI_4
FTI_5
FTI_6
FTI_7
FTI_8
FTI_9
TBG_measured
TBG
referral_source
Class
Test mode: 5-fold cross-validation

==== Classifier model (full training set) ====
J48 pruned tree
-----
T3_1 = '(-inf-1.105)'
| T3_measured = t
| | TT4_1 = '(-inf-44.8)': negative (54.0/6.0)
| | TT4_1 = '(44.8-inf)'
| | | FTI_4 = '(-inf-159.2)'
| | | | referral_source = SVHC: sick (9.0)
| | | | referral_source = other
| | | | | T4U_measured = t
| | | | | T4U_4 = '(-inf-1.078)'
| | | | | | age_1 = '(-inf-46.4]'
| | | | | | | on_thyroxine = f: negative (13.0/1.0)
| | | | | | | on_thyroxine = t: sick (2.0)
| | | | | | | age_1 = '(46.4-inf)'
| | | | | | | TT4_3 = '(-inf-130.4)': sick (27.0/9.0)
| | | | | | | TT4_3 = '(130.4-inf)': negative (2.0)
| | | | | | | T4U_4 = '(1.078-inf)': negative (3.0)
| | | | | | | T4U_measured = f: negative (2.77)
| | | | | | | referral_source = SVI
| | | | | | | FTI_2 = '(-inf-80.6]'
| | | | | | | | age_1 = '(-inf-46.4)': sick (7.3)
| | | | | | | | age_1 = '(46.4-inf)'
| | | | | | | | sick = f
| | | | | | | | T4U_2 = '(-inf-0.664)': negative (2.1/0.1)
| | | | | | | | T4U_2 = '(0.664-inf)'
| | | | | | | | sex = F: negative (10.53/3.53)
| | | | | | | | sex = M: sick (6.26/1.0)
| | | | | | | | sick = t: sick (2.0)
| | | | | | | | FTI_2 = '(80.6-inf)': sick (147.2/2.0)
| | | | | | | | referral_source = STMV: sick (0.0)
| | | | | | | | referral_source = SVHD: sick (3.0)
| | | | | | | | FTI_4 = '(159.2-inf)': negative (19.85/3.62)
| T3_measured = f: negative (79.64/0.31)
T3_1 = '(1.105-inf)': negative (3381.36/24.69)

Number of Leaves : 19
Size of the tree : 34

Time taken to build model: 0.18 seconds

==== Stratified cross-validation ====
==== Summary ====
Correctly Classified Instances      3703          98.1707 %
Incorrectly Classified Instances     69           1.8293 %
Kappa statistic                      0.831
Mean absolute error                  0.0281
Root mean squared error              0.1244
Relative absolute error              24.3465 %
Root relative squared error         51.8784 %
Total Number of Instances            3772

==== Detailed Accuracy By Class ====
      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC    ROC Area   PRC Area   Class
      0.994     0.212     0.986     0.994     0.990     0.833     0.918     0.988   negative
      0.788     0.006     0.901     0.788     0.841     0.833     0.918     0.842     sick
Weighted Avg.      0.982     0.199     0.981     0.982     0.981     0.833     0.918     0.979

==== Confusion Matrix ====
      a      b  <-- classified as
3521    20 |     a = negative
        49  182 |     b = sick

```

## Cross Validation 10 folds

```
==== Run information ====

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B10-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  78
              age_1
              age_2
              age_3
              age_4
              age_5
              age_6
              age_7
              age_8
              age_9
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              TSH_2
              TSH_3
              TSH_4
              TSH_5
              TSH_6
              TSH_7
              TSH_8
              TSH_9
              T3_measured
              T3_1
              T3_2
              T3_3
              T3_4
              T3_5
              T3_6
              T3_7
              T3_8
              T3_9
              TT4_measured
              TT4_1
              TT4_2
              TT4_3
              TT4_4
              TT4_5
              TT4_6
              TT4_7
              TT4_8
              TT4_9
              T4U_measured
              T4U_1
              T4U_2
              T4U_3
              T4U_4
              T4U_5
              T4U_6
              T4U_7
              T4U_8
              T4U_9
              FTI_measured
              FTI_1
              FTI_2
              FTI_3
              FTI_4
              FTI_5
              FTI_6
              FTI_7
              FTI_8
              FTI_9
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   10-fold cross-validation
```

```
==== Classifier model (full training set) ====
J48 pruned tree
-----
T3_1 = '(-inf-1.105'
| T3_measured = t
| | TT4_1 = '(-inf-44.8]': negative (54.0/6.0)
| | TT4_1 = '(44.8-inf)'
| | | FTI_4 = '(-inf-159.2]'
| | | | referral_source = SVHC: sick (9.0)
| | | | referral_source = other
| | | | | T4U_measured = t
| | | | | T4U_4 = '(-inf-1.078]
| | | | | | age_1 = '(-inf-46.4]
| | | | | | | on_thyroxine = f: negative (13.0/1.0)
| | | | | | | on_thyroxine = t: sick (2.0)
| | | | | | | age_1 = '(46.4-inf)'
| | | | | | | TT4_3 = '(-inf-130.4]': sick (27.0/9.0)
| | | | | | | TT4_3 = '(130.4-inf)': negative (2.0)
| | | | | | | T4U_4 = '(1.078-inf)': negative (3.0)
| | | | | | | T4U_measured = f: negative (2.77)
| | | | | | | referral_source = SVI
| | | | | | | FTI_2 = '(-inf-80.6]'
| | | | | | | age_1 = '(-inf-46.4]': sick (7.3)
| | | | | | | age_1 = '(46.4-inf)'
| | | | | | | sick = f
| | | | | | | | T4U_2 = '(-inf-0.664]': negative (2.1/0.1)
| | | | | | | | T4U_2 = '(0.664-inf)'
| | | | | | | | sex = F: negative (10.53/3.53)
| | | | | | | | sex = M: sick (6.26/1.0)
| | | | | | | | sick = t: sick (2.0)
| | | | | | | | FTI_2 = '(80.6-inf)': sick (147.2/2.0)
| | | | | | | | referral_source = STMW: sick (0.0)
| | | | | | | | referral_source = SVHD: sick (3.0)
| | | | | | | | FTI_4 = '(159.2-inf)': negative (19.85/3.62)
| | T3_measured = f: negative (79.64/0.31)
T3_1 = '(1.105-inf)': negative (3381.36/24.69)
```

Number of Leaves : 19

Size of the tree : 34

Time taken to build model: 0.16 seconds

==== Stratified cross-validation ====

==== Summary ====

Correctly Classified Instances	3704	98.1972 %
Incorrectly Classified Instances	68	1.8028 %
Kappa statistic	0.8331	
Mean absolute error	0.0285	
Root mean squared error	0.1254	
Relative absolute error	24.7223 %	
Root relative squared error	52.3065 %	
Total Number of Instances	3772	

==== Detailed Accuracy By Class ====

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
0.995	0.212	0.986	0.995	0.990	0.835	0.915	0.986		negative
0.788	0.005	0.905	0.788	0.843	0.835	0.915	0.825		sick
Weighted Avg.	0.982	0.199	0.981	0.982	0.981	0.835	0.915	0.976	

==== Confusion Matrix ====

a	b	<-- classified as
3522	19	a = negative
49	182	b = sick

## Cross Validation 15 folds

==== Run information ====

```
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B10-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  78
              age_1
              age_2
              age_3
              age_4
```

```
age_5
age_6
age_7
age_8
age_9
sex
on_thyroxine
query_on_thyroxine
on_antithyroid_medication
sick
pregnant
thyroid_surgery
I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH_1
TSH_2
TSH_3
TSH_4
TSH_5
TSH_6
TSH_7
TSH_8
TSH_9
T3_measured
T3_1
T3_2
T3_3
T3_4
T3_5
T3_6
T3_7
T3_8
T3_9
TT4_measured
TT4_1
TT4_2
TT4_3
TT4_4
TT4_5
TT4_6
TT4_7
TT4_8
TT4_9
T4U_measured
T4U_1
T4U_2
T4U_3
T4U_4
T4U_5
T4U_6
T4U_7
T4U_8
T4U_9
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
FTI_5
FTI_6
FTI_7
FTI_8
FTI_9
TBG_measured
TBG
referral_source
Class
Test mode: 15-fold cross-validation
==== Classifier model (full training set) ====
J48 pruned tree
-----
T3_1 = '(-inf-1.105'
|   T3_measured = t
|   |   TT4_1 = '(-inf-44.8]': negative (54.0/6.0)
|   |   TT4_1 = '(44.8-inf)'
|   |   |   FTI_4 = '(-inf-159.2]'
|   |   |   referral_source = SVHC: sick (9.0)
```

```

|   |   |   |   referral_source = other
|   |   |   |   T4U_measured = t
|   |   |   |   |   T4U_4 = '(-inf-1.078]'
|   |   |   |   |   age_1 = '(-inf-46.4)'
|   |   |   |   |   |   on_thyroxine = f: negative (13.0/1.0)
|   |   |   |   |   |   on_thyroxine = t: sick (2.0)
|   |   |   |   |   age_1 = '(46.4-inf)'
|   |   |   |   |   |   TT4_3 = '(-inf-130.4)': sick (27.0/9.0)
|   |   |   |   |   |   TT4_3 = '(130.4-inf)': negative (2.0)
|   |   |   |   |   T4U_4 = '(1.078-inf)': negative (3.0)
|   |   |   |   |   T4U_measured = f: negative (2.77)
|   |   |   |   referral_source = SVI
|   |   |   |   FTI_2 = '(-inf-80.6]'
|   |   |   |   |   age_1 = '(-inf-46.4)': sick (7.3)
|   |   |   |   |   age_1 = '(46.4-inf)'
|   |   |   |   |   |   sick = f
|   |   |   |   |   |   |   T4U_2 = '(-inf-0.664)': negative (2.1/0.1)
|   |   |   |   |   |   T4U_2 = '(0.664-inf)'
|   |   |   |   |   |   sex = F: negative (10.53/3.53)
|   |   |   |   |   |   sex = M: sick (6.26/1.0)
|   |   |   |   |   |   sick = t: sick (2.0)
|   |   |   |   |   FTI_2 = '(80.6-inf)': sick (147.2/2.0)
|   |   |   |   |   referral_source = STMW: sick (0.0)
|   |   |   |   |   referral_source = SVHD: sick (3.0)
|   |   |   |   FTI_4 = '(159.2-inf)': negative (19.85/3.62)
|   T3_measured = f: negative (79.64/0.31)
T3_1 = '(1.105-inf)': negative (3381.36/24.69)

Number of Leaves : 19
Size of the tree : 34

Time taken to build model: 0.15 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3704          98.1972 %
Incorrectly Classified Instances    68           1.8028 %
Kappa statistic                   0.8338
Mean absolute error               0.0283
Root mean squared error          0.1255
Relative absolute error          24.5397 %
Root relative squared error     52.3518 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC    ROC Area  PRC Area  Class
0.994    0.208    0.987    0.994    0.990    0.836    0.906    0.984    negative
0.792    0.006    0.901    0.792    0.843    0.836    0.906    0.799    sick
Weighted Avg.    0.982    0.195    0.981    0.982    0.981    0.836    0.906    0.973

==== Confusion Matrix ====

      a      b  <-- classified as
3521   20 |     a = negative
      48   183 |     b = sick

```

## Percentage Split 50%

```

==== Run information ====

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B10-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  78
              age_1
              age_2
              age_3
              age_4
              age_5
              age_6
              age_7
              age_8
              age_9
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery

```

```

I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH_1
TSH_2
TSH_3
TSH_4
TSH_5
TSH_6
TSH_7
TSH_8
TSH_9
T3_measured
T3_1
T3_2
T3_3
T3_4
T3_5
T3_6
T3_7
T3_8
T3_9
TT4_measured
TT4_1
TT4_2
TT4_3
TT4_4
TT4_5
TT4_6
TT4_7
TT4_8
TT4_9
T4U_measured
T4U_1
T4U_2
T4U_3
T4U_4
T4U_5
T4U_6
T4U_7
T4U_8
T4U_9
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
FTI_5
FTI_6
FTI_7
FTI_8
FTI_9
TBG_measured
TBG
referral_source
Class
Test mode: split 50.0% train, remainder test

==== Classifier model (full training set) ====
J48 pruned tree
-----
T3_1 = '(-inf-1.105'
| T3_measured = t
| | TT4_1 = '(-inf-44.8]': negative (54.0/6.0)
| | TT4_1 = '(44.8-inf)'
| | | FTI_4 = '(-inf-159.2]'
| | | | referral_source = SVHC: sick (9.0)
| | | | referral_source = other
| | | | | T4U_measured = t
| | | | | | T4U_4 = '(-inf-1.078]
| | | | | | | age_1 = '(-inf-46.4]'
| | | | | | | on_thyroxine = f: negative (13.0/1.0)
| | | | | | | on_thyroxine = t: sick (2.0)
| | | | | | | age_1 = '(46.4-inf)'
| | | | | | | TT4_3 = '(-inf-130.4]': sick (27.0/9.0)
| | | | | | | TT4_3 = '(130.4-inf)': negative (2.0)
| | | | | | | T4U_4 = '(1.078-inf)': negative (3.0)
| | | | | | | T4U_measured = f: negative (2.77)
| | | | | | | referral_source = SVI

```

```

|   |   |   |   |   FTI_2 = '(-inf-80.6]'
|   |   |   |   |   age_1 = '(-inf-46.4]': sick (7.3)
|   |   |   |   |   age_1 = '(46.4-inf)'
|   |   |   |   |   sick = f
|   |   |   |   |   |   T4U_2 = '(-inf-0.664]': negative (2.1/0.1)
|   |   |   |   |   |   T4U_2 = '(0.664-inf)'
|   |   |   |   |   |   sex = F: negative (10.53/3.53)
|   |   |   |   |   |   sex = M: sick (6.26/1.0)
|   |   |   |   |   |   sick = t: sick (2.0)
|   |   |   |   |   FTI_2 = '(80.6-inf)': sick (147.2/2.0)
|   |   |   |   referral_source = STMW: sick (0.0)
|   |   |   |   referral_source = SVHD: sick (3.0)
|   |   |   FTI_4 = '(159.2-inf)': negative (19.85/3.62)
|   T3_measured = f: negative (79.64/0.31)
T3_1 = '(1.105-inf)': negative (3381.36/24.69)

Number of Leaves : 19
Size of the tree : 34

Time taken to build model: 0.17 seconds
== Evaluation on test split ==
Time taken to test model on test split: 0.04 seconds
== Summary ==
Correctly Classified Instances      1853          98.2503 %
Incorrectly Classified Instances    33           1.7497 %
Kappa statistic                      0.8315
Mean absolute error                  0.0265
Root mean squared error              0.1221
Relative absolute error              23.0089 %
Root relative squared error         50.6306 %
Total Number of Instances           1886

== Detailed Accuracy By Class ==


|               | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC   | ROC Area | PRC Area | Class |
|---------------|---------|---------|-----------|--------|-----------|-------|----------|----------|-------|
| 0.998         | 0.256   | 0.983   | 0.998     | 0.991  | 0.840     | 0.945 | 0.993    | negative |       |
| 0.744         | 0.002   | 0.967   | 0.744     | 0.841  | 0.840     | 0.945 | 0.841    | sick     |       |
| Weighted Avg. | 0.983   | 0.241   | 0.982     | 0.983  | 0.981     | 0.840 | 0.945    | 0.984    |       |


== Confusion Matrix ==


| a    | b  | <-- classified as |
|------|----|-------------------|
| 1766 | 3  | a = negative      |
| 30   | 87 | b = sick          |


```

## Percentage Split 60%

```

== Run information ==
Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2
Relation: sick-weka.filters.unsupervised.attribute.Discretize-D-B10-M-1.0-Rfirst-last-precision6
Instances: 3772
Attributes: 78
age_1
age_2
age_3
age_4
age_5
age_6
age_7
age_8
age_9
sex
on_thyroxine
query_on_thyroxine
on_antithyroid_medication
sick
pregnant
thyroid_surgery
I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured

```

```

TSH_1
TSH_2
TSH_3
TSH_4
TSH_5
TSH_6
TSH_7
TSH_8
TSH_9
T3_measured
T3_1
T3_2
T3_3
T3_4
T3_5
T3_6
T3_7
T3_8
T3_9
TT4_measured
TT4_1
TT4_2
TT4_3
TT4_4
TT4_5
TT4_6
TT4_7
TT4_8
TT4_9
T4U_measured
T4U_1
T4U_2
T4U_3
T4U_4
T4U_5
T4U_6
T4U_7
T4U_8
T4U_9
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
FTI_5
FTI_6
FTI_7
FTI_8
FTI_9
TBG_measured
TBG
referral_source
Class
Test mode: split 60.0% train, remainder test

==== Classifier model (full training set) ====
J48 pruned tree
-----
T3_1 = '(-inf-1.105]'
| T3_measured = t
| | TT4_1 = '(-inf-44.8]': negative (54.0/6.0)
| | TT4_1 = '(44.8-inf)'
| | | FTI_4 = '(-inf-159.2]'
| | | | referral_source = SVHC: sick (9.0)
| | | | referral_source = other
| | | | T4U_measured = t
| | | | | T4U_4 = '(-inf-1.078]'
| | | | | | age_1 = '(-inf-46.4)'
| | | | | | | on_thyroxine = f: negative (13.0/1.0)
| | | | | | | on_thyroxine = t: sick (2.0)
| | | | | | | age_1 = '(46.4-inf)'
| | | | | | | | TT4_3 = '(-inf-130.4)': sick (27.0/9.0)
| | | | | | | | TT4_3 = '(130.4-inf)': negative (2.0)
| | | | | | | T4U_4 = '(1.078-inf)': negative (3.0)
| | | | | | | T4U_measured = f: negative (2.77)
| | | | | | | referral_source = SVI
| | | | | | | FTI_2 = '(-inf-80.6]'
| | | | | | | age_1 = '(-inf-46.4)': sick (7.3)
| | | | | | | age_1 = '(46.4-inf)'
| | | | | | | | sick = f
| | | | | | | | T4U_2 = '(-inf-0.664)': negative (2.1/0.1)
| | | | | | | | T4U_2 = '(0.664-inf)'
| | | | | | | | sex = F: negative (10.53/3.53)
| | | | | | | | sex = M: sick (6.26/1.0)
| | | | | | | | sick = t: sick (2.0)

```

```

|   |   |   |   FTI_2 = '(80.6-inf)': sick (147.2/2.0)
|   |   |   referral_source = STMW: sick (0.0)
|   |   |   referral_source = SVHD: sick (3.0)
|   |   |   FTI_4 = '(159.2-inf)': negative (19.85/3.62)
|   T3_measured = f: negative (79.64/0.31)
T3_1 = '(1.105-inf)': negative (3381.36/24.69)

Number of Leaves : 19

Size of the tree : 34

Time taken to build model: 0.13 seconds

==== Evaluation on test split ====

Time taken to test model on test split: 0 seconds

==== Summary ====

Correctly Classified Instances      1481          98.1445 %
Incorrectly Classified Instances    28           1.8555 %
Kappa statistic                   0.8087
Mean absolute error               0.0279
Root mean squared error          0.1248
Relative absolute error           24.3117 %
Root relative squared error      52.9649 %
Total Number of Instances        1509

==== Detailed Accuracy By Class ====

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC     ROC Area  PRC Area  Class
      0.999    0.292    0.982     0.999    0.990     0.820    0.942    0.993    negative
      0.708    0.001    0.969     0.708    0.818     0.820    0.942    0.833    sick
Weighted Avg.      0.981    0.275    0.981     0.981    0.980     0.820    0.942    0.983

==== Confusion Matrix ====

      a      b  <-- classified as
1418    2 |      a = negative
      26   63 |      b = sick

```

## Percentage Split 66%

```

==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B10-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  78
              age_1
              age_2
              age_3
              age_4
              age_5
              age_6
              age_7
              age_8
              age_9
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              TSH_2
              TSH_3
              TSH_4
              TSH_5
              TSH_6
              TSH_7
              TSH_8
              TSH_9

```

```

T3_measured
T3_1
T3_2
T3_3
T3_4
T3_5
T3_6
T3_7
T3_8
T3_9
TT4_measured
TT4_1
TT4_2
TT4_3
TT4_4
TT4_5
TT4_6
TT4_7
TT4_8
TT4_9
T4U_measured
T4U_1
T4U_2
T4U_3
T4U_4
T4U_5
T4U_6
T4U_7
T4U_8
T4U_9
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
FTI_5
FTI_6
FTI_7
FTI_8
FTI_9
TBG_measured
TBG
referral_source
Class
Test mode: split 66.0% train, remainder test

==== Classifier model (full training set) ====

J48 pruned tree
-----
T3_1 = '(-inf-1.105]'
| T3_measured = t
| | TT4_1 = '(-inf-44.8]': negative (54.0/6.0)
| | TT4_1 = '(44.8-inf)'
| | | FTI_4 = '(-inf-159.2]'
| | | | referral_source = SVHC: sick (9.0)
| | | | referral_source = other
| | | | T4U_measured = t
| | | | | T4U_4 = '(-inf-1.078]'
| | | | | | age_1 = '(-inf-46.4]'
| | | | | | on_thyroxine = f: negative (13.0/1.0)
| | | | | | on_thyroxine = t: sick (2.0)
| | | | | | age_1 = '(46.4-inf)'
| | | | | | TT4_3 = '(-inf-130.4)': sick (27.0/9.0)
| | | | | | TT4_3 = '(130.4-inf)': negative (2.0)
| | | | | | T4U_4 = '(1.078-inf)': negative (3.0)
| | | | | | T4U_measured = f: negative (2.77)
| | | | | | referral_source = SVI
| | | | | | FTI_2 = '(-inf-80.6]'
| | | | | | | age_1 = '(-inf-46.4]': sick (7.3)
| | | | | | | age_1 = '(46.4-inf)'
| | | | | | | sick = f
| | | | | | | T4U_2 = '(-inf-0.664)': negative (2.1/0.1)
| | | | | | | T4U_2 = '(0.664-inf)'
| | | | | | | sex = F: negative (10.53/3.53)
| | | | | | | sex = M: sick (6.26/1.0)
| | | | | | | sick = t: sick (2.0)
| | | | | | | FTI_2 = '(80.6-inf)': sick (147.2/2.0)
| | | | | | | referral_source = STMW: sick (0.0)
| | | | | | | referral_source = SVHD: sick (3.0)
| | | | | | | FTI_4 = '(159.2-inf)': negative (19.85/3.62)
| | T3_measured = f: negative (79.64/0.31)
T3_1 = '(1.105-inf)': negative (3381.36/24.69)

```

Number of Leaves : 19

```

Size of the tree : 34

Time taken to build model: 0.14 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0.02 seconds

==== Summary ===

    Correctly Classified Instances      1249          97.4259 %
    Incorrectly Classified Instances   33           2.5741 %
    Kappa statistic                   0.7189
    Mean absolute error              0.033
    Root mean squared error          0.1351
    Relative absolute error          28.9527 %
    Root relative squared error     57.5694 %
    Total Number of Instances        1282

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.998     0.400     0.976     0.998     0.986     0.739     0.940     0.993     negative
      0.600     0.002     0.938     0.600     0.732     0.739     0.940     0.813     sick
Weighted Avg.   0.974     0.377     0.973     0.974     0.972     0.739     0.940     0.982

==== Confusion Matrix ===

      a     b  <-- classified as
1204   3 |   a = negative
      30   45 |   b = sick

```

## Leave one out fold

```

==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B10-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  78
             age_1
             age_2
             age_3
             age_4
             age_5
             age_6
             age_7
             age_8
             age_9
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             i131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary
             psych
             TSH_measured
             TSH_1
             TSH_2
             TSH_3
             TSH_4
             TSH_5
             TSH_6
             TSH_7
             TSH_8
             TSH_9
             T3_measured
             T3_1
             T3_2
             T3_3
             T3_4
             T3_5
             T3_6
             T3_7
             T3_8

```

```

T3_9
TT4_measured
TT4_1
TT4_2
TT4_3
TT4_4
TT4_5
TT4_6
TT4_7
TT4_8
TT4_9
T4U_measured
T4U_1
T4U_2
T4U_3
T4U_4
T4U_5
T4U_6
T4U_7
T4U_8
T4U_9
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
FTI_5
FTI_6
FTI_7
FTI_8
FTI_9
TBG_measured
TBG
referral_source
Class
Test mode: 3772-fold cross-validation

==== Classifier model (full training set) ===

J48 pruned tree
-----
T3_1 = '(-inf-1.105'
| T3_measured = t
| | TT4_1 = '(-inf-44.8)': negative (54.0/6.0)
| | TT4_1 = '(44.8-inf'
| | | FTI_4 = '(-inf-159.2'
| | | | referral_source = SVHC: sick (9.0)
| | | | referral_source = other
| | | | T4U_measured = t
| | | | | T4U_4 = '(-inf-1.078'
| | | | | | age_1 = '(-inf-46.4'
| | | | | | | on_thyroxine = f: negative (13.0/1.0)
| | | | | | | on_thyroxine = t: sick (2.0)
| | | | | | | age_1 = '(46.4-inf'
| | | | | | | | TT4_3 = '(-inf-130.4': sick (27.0/9.0)
| | | | | | | | TT4_3 = '(130.4-inf)': negative (2.0)
| | | | | | | | T4U_4 = '(1.078-inf)': negative (3.0)
| | | | | | | | T4U_measured = f: negative (2.77)
| | | | | | | | referral_source = SVI
| | | | | | | | FTI_2 = '(-inf-80.6'
| | | | | | | | | age_1 = '(-inf-46.4': sick (7.3)
| | | | | | | | | age_1 = '(46.4-inf'
| | | | | | | | | sick = f
| | | | | | | | | | T4U_2 = '(-inf-0.664)': negative (2.1/0.1)
| | | | | | | | | | T4U_2 = '(0.664-inf'
| | | | | | | | | | sex = F: negative (10.53/3.53)
| | | | | | | | | | sex = M: sick (6.26/1.0)
| | | | | | | | | | sick = t: sick (2.0)
| | | | | | | | | | FTI_2 = '(80.6-inf)': sick (147.2/2.0)
| | | | | | | | | | referral_source = STMW: sick (0.0)
| | | | | | | | | | referral_source = SVHD: sick (3.0)
| | | | | | | | | | FTI_4 = '(159.2-inf)': negative (19.85/3.62)
| | T3_measured = f: negative (79.64/0.31)
T3_1 = '(1.105-inf)': negative (3381.36/24.69)

Number of Leaves : 19
Size of the tree : 34

Time taken to build model: 0.14 seconds

==== Stratified cross-validation ===
==== Summary ===

Correctly Classified Instances      3698          98.0382 %

```

```

Incorrectly Classified Instances      74          1.9618 %
Kappa statistic                      0.8207
Mean absolute error                  0.0288
Root mean squared error              0.1286
Relative absolute error              25.0368 %
Root relative squared error         53.6012 %
Total Number of Instances           3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC     ROC Area   PRC Area   Class
0.993      0.212      0.986      0.993      0.990      0.822      0.897      0.985    negative
0.788      0.007      0.879      0.788      0.831      0.822      0.897      0.841    sick
Weighted Avg.   0.980      0.200      0.980      0.980      0.980      0.822      0.897      0.977

==== Confusion Matrix ====
a      b  <-- classified as
3516  25 |  a = negative
 49  182 |  b = sick

```

#### 4.1.4 40 bins

##### Cross Validation 5 folds

```

==== Run information ====
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B40-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  258
              [list of attributes omitted]
Test mode:   5-fold cross-validation

==== Classifier model (full training set) ====
J48 pruned tree
-----
T3_4 = '(-inf-1.105]'
| T3_measured = t
| | TT4_3 = '(-inf-34.1)': negative (42.0/1.0)
| | TT4_3 = '(34.1-inf)'
| | | TT4_5 = '(-inf-55.5)'
| | | | TSH_2 = '(-inf-26.50475)'
| | | | FTI_8 = '(-inf-80.6)'
| | | | | age_4 = '(-inf-46.4)': sick (5.0)
| | | | | age_4 = '(46.4-inf)'
| | | | | sex = F: negative (10.0/1.0)
| | | | | sex = M: sick (6.0/1.0)
| | | | | FTI_8 = '(80.6-inf)': negative (7.0)
| | | | | TSH_2 = '(26.50475-inf)': negative (7.0)
| | | | TT4_5 = '(55.5-inf)'
| | | | | referral_source = SVHC: sick (9.0)
| | | | | referral_source = other
| | | | | sick = f
| | | | | T4U_measured = t
| | | | | | T4U_6 = '(-inf-0.5605)': sick (3.0)
| | | | | | T4U_6 = '(0.5605-inf)'
| | | | | | | T4U_9 = '(-inf-0.71575)': negative (3.0)
| | | | | | | T4U_9 = '(0.71575-inf)'
| | | | | | | age_6 = '(-inf-69.1]'
| | | | | | | | TSH_1 = '(-inf-13.254875)': negative (16.73)
| | | | | | | | TSH_1 = '(13.254875-inf)': sick (6.27/1.27)
| | | | | | | | age_6 = '(69.1-inf)'
| | | | | | | | on_thyroxine = f: sick (13.0/2.0)
| | | | | | | | on_thyroxine = t: negative (2.0)
| | | | | | | | T4U_measured = f: negative (3.0)
| | | | | | | | sick = t: sick (3.0)
| | | | | | | | referral_source = SVI
| | | | | | | | FTI_16 = '(-inf-159.2]'
| | | | | | | | FTI_6 = '(-inf-60.95)': negative (2.1/0.1)
| | | | | | | | FTI_6 = '(60.95-inf)': sick (157.32)
| | | | | | | | FTI_16 = '(159.2-inf)'
| | | | | | | | TT4_9 = '(-inf-98.3)': sick (2.15)
| | | | | | | | TT4_9 = '(98.3-inf)': negative (10.44/0.44)
| | | | | | | | referral_source = STMW: sick (0.0)
| | | | | | | | referral_source = SVHD: sick (3.0)
| | T3_measured = f: negative (79.64/0.31)
T3_4 = '(1.105-inf)'
| FTI_17 = '(-inf-169.025)': negative (3222.3/8.69)
| FTI_17 = '(169.025-inf)'
| | T3_10 = '(-inf-2.6875)'
```

```

|   |   |   on_thyroxine = f
|   |   |   T3_measured = t
|   |   |   |   TT4_14 = '(-inf-151.8]': negative (11.57)
|   |   |   |   TT4_14 = '(151.8-inf)'
|   |   |   |   |   TT4_18 = '(-inf-194.6]': sick (20.78/4.78)
|   |   |   |   |   TT4_18 = '(194.6-inf)': negative (3.09)
|   |   |   |   T3_measured = f: negative (11.21)
|   |   |   on_thyroxine = t: negative (25.63)
|   |   T3_10 = '(2.6875-inf)': negative (86.77)

Number of Leaves : 29
Size of the tree : 54

Time taken to build model: 0.36 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3713          98.4358 %
Incorrectly Classified Instances    59           1.5642 %
Kappa statistic                   0.8609
Mean absolute error               0.02
Root mean squared error          0.1167
Relative absolute error          17.3566 %
Root relative squared error     48.6876 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
      0.993    0.152    0.990     0.993    0.992     0.861    0.945     0.990    negative
      0.848    0.007    0.891     0.848    0.869     0.861    0.945     0.863    sick
Weighted Avg.      0.984    0.143    0.984     0.984    0.984     0.861    0.945     0.982

==== Confusion Matrix ====

      a     b  <-- classified as
3517  24 |     a = negative
      35 196 |     b = sick

```

## Cross Validation 10 folds

```

==== Run information ====
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B40-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  258
              [list of attributes omitted]
Test mode:   10-fold cross-validation

==== Classifier model (full training set) ====
J48 pruned tree
-----
T3_4 = '(-inf-1.105'
|   T3_measured = t
|   |   TT4_3 = '(-inf-34.1]': negative (42.0/1.0)
|   |   TT4_3 = '(34.1-inf)'
|   |   |   TT4_5 = '(-inf-55.5]'
|   |   |   |   TSH_2 = '(-inf-26.50475]'
|   |   |   |   |   FTI_8 = '(-inf-80.6]'
|   |   |   |   |   age_4 = '(-inf-46.4]': sick (5.0)
|   |   |   |   |   age_4 = '(46.4-inf)'
|   |   |   |   |   sex = F: negative (10.0/1.0)
|   |   |   |   |   sex = M: sick (6.0/1.0)
|   |   |   |   |   FTI_8 = '(80.6-inf)': negative (7.0)
|   |   |   |   |   TSH_2 = '(26.50475-inf)': negative (7.0)
|   |   |   TT4_5 = '(55.5-inf)'
|   |   |   |   referral_source = SVHC: sick (9.0)
|   |   |   |   referral_source = other
|   |   |   |   sick = f
|   |   |   |   |   T4U_measured = t
|   |   |   |   |   |   T4U_6 = '(-inf-0.5605]': sick (3.0)
|   |   |   |   |   |   T4U_6 = '(0.5605-inf)'
|   |   |   |   |   |   T4U_9 = '(-inf-0.71575]': negative (3.0)
|   |   |   |   |   |   T4U_9 = '(0.71575-inf)'
|   |   |   |   |   |   age_6 = '(-inf-69.1]'
|   |   |   |   |   |   |   TSH_1 = '(-inf-13.254875]': negative (16.73)
|   |   |   |   |   |   |   TSH_1 = '(13.254875-inf)': sick (6.27/1.27)
|   |   |   |   |   |   |   age_6 = '(69.1-inf)'

```

```

| | | | | | | | | | on_thyroxine = f: sick (13.0/2.0)
| | | | | | | | | | on_thyroxine = t: negative (2.0)
| | | | | | T4U_measured = f: negative (3.0)
| | | | | | sick = t: sick (3.0)
| | | | | | referral_source = SVI
| | | | | | FTI_16 = '(-inf-159.2)'
| | | | | | | | FTI_6 = '(-inf-60.95)': negative (2.1/0.1)
| | | | | | FTI_6 = '(60.95-inf)': sick (157.32)
| | | | | | FTI_16 = '(159.2-inf)'
| | | | | | TT4_9 = '(-inf-98.3)': sick (2.15)
| | | | | | TT4_9 = '(98.3-inf)': negative (10.44/0.44)
| | | | | | referral_source = STMW: sick (0.0)
| | | | | | referral_source = SVHD: sick (3.0)
| | | | T3_measured = f: negative (79.64/0.31)
T3_4 = '(1.105-inf)'
| | FTI_17 = '(-inf-169.025)': negative (3222.3/8.69)
| | FTI_17 = '(169.025-inf)'
| | | | T3_10 = '(-inf-2.6875)'
| | | | | | on_thyroxine = f
| | | | | | T3_measured = t
| | | | | | TT4_14 = '(-inf-151.8)': negative (11.57)
| | | | | | TT4_14 = '(151.8-inf)'
| | | | | | TT4_18 = '(-inf-194.6)': sick (20.78/4.78)
| | | | | | TT4_18 = '(194.6-inf)': negative (3.09)
| | | | | | T3_measured = f: negative (11.21)
| | | | | | on_thyroxine = t: negative (25.63)
| | | | T3_10 = '(2.6875-inf)': negative (86.77)

Number of Leaves : 29

Size of the tree : 54

Time taken to build model: 0.36 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3721          98.6479 %
Incorrectly Classified Instances   51           1.3521 %
Kappa statistic                   0.8817
Mean absolute error               0.0177
Root mean squared error          0.1101
Relative absolute error          15.397 %
Root relative squared error     45.9352 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ====

          TP Rate  FP Rate  Precision  Recall   F-Measure  MCC    ROC Area  PRC Area  Class
0         0.993   0.117    0.992     0.993    0.993    0.882   0.941    0.989    negative
1         0.883   0.007    0.895     0.883    0.889    0.882   0.941    0.874    sick
Weighted Avg.      0.986   0.110    0.986     0.986    0.986    0.882   0.941    0.982

==== Confusion Matrix ====

      a     b  <-- classified as
3517  24 |     a = negative
      27  204 |     b = sick

```

## Cross Validation 15 folds

```

==== Run information ====
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B40-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  258
              [list of attributes omitted]
Test mode:   15-fold cross-validation

==== Classifier model (full training set) ====

J48 pruned tree
-----
T3_4 = '(-inf-1.105)'
| T3_measured = t
| | TT4_3 = '(-inf-34.1)': negative (42.0/1.0)
| | TT4_3 = '(34.1-inf)'
| | | TT4_5 = '(-inf-55.5)'
| | | TSH_2 = '(-inf-26.50475)'
| | | | FTI_8 = '(-inf-80.6)'
| | | | | aqe_4 = '(-inf-46.4)': sick (5.0)

```

```

| | | | | age_4 = '(46.4-inf)'
| | | | | sex = F: negative (10.0/1.0)
| | | | | sex = M: sick (6.0/1.0)
| | | | | FTI_8 = '(80.6-inf)': negative (7.0)
| | | | | TSH_2 = '(26.50475-inf)': negative (7.0)
| | | | | TT4_5 = '(55.5-inf)'
| | | | | referral_source = SVHC: sick (9.0)
| | | | | referral_source = other
| | | | | sick = f
| | | | | T4U_measured = t
| | | | | | T4U_6 = '(-inf-0.5605)': sick (3.0)
| | | | | | T4U_6 = '(0.5605-inf)'
| | | | | | | T4U_9 = '(-inf-0.71575)': negative (3.0)
| | | | | | | T4U_9 = '(0.71575-inf)'
| | | | | | | age_6 = '(-inf-69.1]'
| | | | | | | | TSH_1 = '(-inf-13.254875)': negative (16.73)
| | | | | | | | TSH_1 = '(13.254875-inf)': sick (6.27/1.27)
| | | | | | | | age_6 = '(69.1-inf)'
| | | | | | | | on_thyroxine = f: sick (13.0/2.0)
| | | | | | | | on_thyroxine = t: negative (2.0)
| | | | | | | | T4U_measured = f: negative (3.0)
| | | | | | | | sick = t: sick (3.0)
| | | | | | | | referral_source = SVI
| | | | | | | | FTI_16 = '(-inf-159.2]'
| | | | | | | | FTI_6 = '(-inf-60.95)': negative (2.1/0.1)
| | | | | | | | FTI_6 = '(60.95-inf)': sick (157.32)
| | | | | | | | FTI_16 = '(159.2-inf)'
| | | | | | | | TT4_9 = '(-inf-98.3)': sick (2.15)
| | | | | | | | TT4_9 = '(98.3-inf)': negative (10.44/0.44)
| | | | | | | | referral_source = STMW: sick (0.0)
| | | | | | | | referral_source = SVHD: sick (3.0)
| | | | | | | | T3_measured = f: negative (79.64/0.31)
T3_4 = '(1.105-inf'
| | FTI_17 = '(-inf-169.025)': negative (3222.3/8.69)
| | FTI_17 = '(169.025-inf)'
| | | T3_10 = '(-inf-2.6875)'
| | | | on_thyroxine = f
| | | | T3_measured = t
| | | | | TT4_14 = '(-inf-151.8)': negative (11.57)
| | | | | TT4_14 = '(151.8-inf)'
| | | | | | TT4_18 = '(-inf-194.6)': sick (20.78/4.78)
| | | | | | TT4_18 = '(194.6-inf)': negative (3.09)
| | | | | | T3_measured = f: negative (11.21)
| | | | | | on_thyroxine = t: negative (25.63)
| | | | | | T3_10 = '(2.6875-inf)': negative (86.77)

Number of Leaves : 29
Size of the tree : 54

Time taken to build model: 0.34 seconds

==== Stratified cross-validation ====
==== Summary ====

    Correctly Classified Instances      3731          98.913 %
    Incorrectly Classified Instances     41           1.087 %
    Kappa statistic                      0.9049
    Mean absolute error                  0.0156
    Root mean squared error              0.0982
    Relative absolute error              13.5431 %
    Root relative squared error         40.9417 %
    Total Number of Instances            3772

==== Detailed Accuracy By Class ====

    TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
    0.995     0.095     0.994      0.995     0.994      0.905     0.943     0.989     negative
    0.905     0.005     0.917      0.905     0.911      0.905     0.943     0.918     sick
    Weighted Avg.   0.989     0.090     0.989      0.989     0.989      0.905     0.943     0.985

==== Confusion Matrix ====
    a     b  <-- classified as
3522  19 |    a = negative
    22  209 |    b = sick

```

## Percentage Split 50%

```

==== Run information ====
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B40-M-1.0-Rfirst-last-precision6

```

```

Instances:      3772
Attributes:    258
               [list of attributes omitted]
Test mode:     split 50.0% train, remainder test

==== Classifier model (full training set) ===

J48 pruned tree
-----

T3_4 = '(-inf-1.105)'
| T3_measured = t
| | TT4_3 = '(-inf-34.1)': negative (42.0/1.0)
| | TT4_3 = '(34.1-inf)'
| | | TT4_5 = '(-inf-55.5)'
| | | | TSH_2 = '(-inf-26.50475)'
| | | | | FTI_8 = '(-inf-80.6)'
| | | | | age_4 = '(-inf-46.4)': sick (5.0)
| | | | | age_4 = '(46.4-inf)'
| | | | | sex = F: negative (10.0/1.0)
| | | | | sex = M: sick (6.0/1.0)
| | | | | FTI_8 = '(80.6-inf)': negative (7.0)
| | | | | TSH_2 = '(26.50475-inf)': negative (7.0)
| | | | TT4_5 = '(55.5-inf)'
| | | | | referral_source = SVHC: sick (9.0)
| | | | | referral_source = other
| | | | | sick = f
| | | | | T4U_measured = t
| | | | | | T4U_6 = '(-inf-0.5605)': sick (3.0)
| | | | | | T4U_6 = '(0.5605-inf)'
| | | | | | T4U_9 = '(-inf-0.71575)': negative (3.0)
| | | | | | T4U_9 = '(0.71575-inf)'
| | | | | | | age_6 = '(-inf-69.1)'
| | | | | | | TSH_1 = '(-inf-13.254875)': negative (16.73)
| | | | | | | TSH_1 = '(13.254875-inf)': sick (6.27/1.27)
| | | | | | | age_6 = '(69.1-inf)'
| | | | | | | on_thyroxine = f: sick (13.0/2.0)
| | | | | | | on_thyroxine = t: negative (2.0)
| | | | | | | T4U_measured = f: negative (3.0)
| | | | | | | sick = t: sick (3.0)
| | | | | | | referral_source = SVI
| | | | | | | FTI_16 = '(-inf-159.2)'
| | | | | | | FTI_6 = '(-inf-60.95)': negative (2.1/0.1)
| | | | | | | FTI_6 = '(60.95-inf)': sick (157.32)
| | | | | | | FTI_16 = '(159.2-inf)'
| | | | | | | TT4_9 = '(-inf-98.3)': sick (2.15)
| | | | | | | TT4_9 = '(98.3-inf)': negative (10.44/0.44)
| | | | | | | referral_source = STMW: sick (0.0)
| | | | | | | referral_source = SVHD: sick (3.0)
| | | | | | T3_measured = f: negative (79.64/0.31)
T3_4 = '(1.105-inf'
| FTI_17 = '(-inf-169.025)': negative (3222.3/8.69)
| FTI_17 = '(169.025-inf'
| | T3_10 = '(-inf-2.6875)'
| | | on_thyroxine = f
| | | | T3_measured = t
| | | | | TT4_14 = '(-inf-151.8)': negative (11.57)
| | | | | TT4_14 = '(151.8-inf)'
| | | | | TT4_18 = '(-inf-194.6)': sick (20.78/4.78)
| | | | | TT4_18 = '(194.6-inf)': negative (3.09)
| | | | | T3_measured = f: negative (11.21)
| | | | | on_thyroxine = t: negative (25.63)
| | | | | T3_10 = '(2.6875-inf)': negative (86.77)

Number of Leaves : 29

Size of the tree : 54

Time taken to build model: 0.39 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0.04 seconds

==== Summary ===

Correctly Classified Instances      1853          98.2503 %
Incorrectly Classified Instances   33           1.7497 %
Kappa statistic                   0.8427
Mean absolute error               0.026
Root mean squared error          0.1277
Relative absolute error          22.5557 %
Root relative squared error     52.9324 %
Total Number of Instances        1886

==== Detailed Accuracy By Class ===

```

```

TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
0.994   0.188    0.988     0.994   0.991    0.844    0.894     0.985    negative
0.812   0.006    0.896     0.812   0.852    0.844    0.894     0.771    sick
Weighted Avg. 0.983   0.177    0.982     0.983   0.982    0.844    0.894     0.972

==== Confusion Matrix ====
a      b      <-- classified as
1758   11 |      a = negative
22     95 |      b = sick

```

## Percentage Split 60%

```

==== Run information ====
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B40-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  258
              [list of attributes omitted]
Test mode:   split 60.0% train, remainder test

==== Classifier model (full training set) ====
J48 pruned tree
-----
T3_4 = '(-inf-1.105]'
| T3_measured = t
| | TT4_3 = '(-inf-34.1)': negative (42.0/1.0)
| | TT4_3 = '(34.1-inf)'
| | | TT4_5 = '(-inf-55.5)'
| | | | TSH_2 = '(-inf-26.50475)'
| | | | | FTI_8 = '(-inf-80.6)'
| | | | | age_4 = '(-inf-46.4)': sick (5.0)
| | | | | age_4 = '(46.4-inf)'
| | | | | sex = F: negative (10.0/1.0)
| | | | | sex = M: sick (6.0/1.0)
| | | | | FTI_8 = '(80.6-inf)': negative (7.0)
| | | | | TSH_2 = '(26.50475-inf)': negative (7.0)
| | | | | TT4_5 = '(55.5-inf)'
| | | | | referral_source = SVHC: sick (9.0)
| | | | | referral_source = other
| | | | | sick = f
| | | | | T4U_measured = t
| | | | | | T4U_6 = '(-inf-0.5605)': sick (3.0)
| | | | | | T4U_6 = '(0.5605-inf)'
| | | | | | T4U_9 = '(-inf-0.71575)': negative (3.0)
| | | | | | T4U_9 = '(0.71575-inf)'
| | | | | | age_6 = '(-inf-69.1)'
| | | | | | | TSH_1 = '(-inf-13.254875)': negative (16.73)
| | | | | | | TSH_1 = '(13.254875-inf)': sick (6.27/1.27)
| | | | | | | age_6 = '(69.1-inf)'
| | | | | | | on_thyroxine = f: sick (13.0/2.0)
| | | | | | | on_thyroxine = t: negative (2.0)
| | | | | | | T4U_measured = f: negative (3.0)
| | | | | | sick = t: sick (3.0)
| | | | | | referral_source = SVI
| | | | | | FTI_16 = '(-inf-159.2)'
| | | | | | FTI_6 = '(-inf-60.95)': negative (2.1/0.1)
| | | | | | FTI_6 = '(60.95-inf)': sick (157.32)
| | | | | | FTI_16 = '(159.2-inf)'
| | | | | | TT4_9 = '(-inf-98.3)': sick (2.15)
| | | | | | TT4_9 = '(98.3-inf)': negative (10.44/0.44)
| | | | | | referral_source = STMW: sick (0.0)
| | | | | | referral_source = SVHD: sick (3.0)
| | T3_measured = f: negative (79.64/0.31)
T3_4 = '(1.105-inf)'
| FTI_17 = '(-inf-169.025)': negative (3222.3/8.69)
| FTI_17 = '(169.025-inf)'
| | T3_10 = '(-inf-2.6875)'
| | | on_thyroxine = f
| | | | T3_measured = t
| | | | | TT4_14 = '(-inf-151.8)': negative (11.57)
| | | | | TT4_14 = '(151.8-inf)'
| | | | | TT4_18 = '(-inf-194.6)': sick (20.78/4.78)
| | | | | TT4_18 = '(194.6-inf)': negative (3.09)
| | | | | T3_measured = f: negative (11.21)
| | | | | on_thyroxine = t: negative (25.63)
| | | | | T3_10 = '(2.6875-inf)': negative (86.77)

```

Number of Leaves : 29

```

Size of the tree : 54

Time taken to build model: 0.52 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0.07 seconds

==== Summary ===

Correctly Classified Instances      1489          98.6746 %
Incorrectly Classified Instances    20           1.3254 %
Kappa statistic                      0.8711
Mean absolute error                  0.0213
Root mean squared error              0.1127
Relative absolute error              18.5329 %
Root relative squared error         47.8443 %
Total Number of Instances           1509

==== Detailed Accuracy By Class ===

          TP Rate   FP Rate   Precision   Recall   F-Measure   MCC   ROC Area   PRC Area   Class
          0.998     0.191     0.988     0.998     0.993     0.875     0.888     0.985   negative
          0.809     0.002     0.960     0.809     0.878     0.875     0.888     0.836   sick
Weighted Avg.       0.987     0.180     0.986     0.987     0.986     0.875     0.888     0.976

==== Confusion Matrix ===

      a      b    <-- classified as
1417    3 |    a = negative
      17   72 |    b = sick

```

## Percentage Split 66%

```

==== Run information ====
Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2
Relation: sick-weka.filters.unsupervised.attribute.Discretize-D-B40-M-1.0-Rfirst-last-precision6
Instances: 3772
Attributes: 258
[ list of attributes omitted]
Test mode: split 66.0% train, remainder test

==== Classifier model (full training set) ====
-----


T3_4 = '(-inf-1.105]'
| T3_measured = t
| | TT4_3 = '(-inf-34.1]': negative (42.0/1.0)
| | TT4_3 = '(34.1-inf)'
| | | TT4_5 = '(-inf-55.5]'
| | | | TSH_2 = '(-inf-26.50475]'
| | | | | FTI_8 = '(-inf-80.6]'
| | | | | age_4 = '(-inf-46.4]': sick (5.0)
| | | | | age_4 = '(46.4-inf)'
| | | | | sex = F: negative (10.0/1.0)
| | | | | sex = M: sick (6.0/1.0)
| | | | | FTI_8 = '(80.6-inf)': negative (7.0)
| | | | | TSH_2 = '(26.50475-inf)': negative (7.0)
| | | TT4_5 = '(55.5-inf'
| | | | referral_source = SVHC: sick (9.0)
| | | | referral_source = other
| | | | sick = f
| | | | | T4U_measured = t
| | | | | | T4U_6 = '(-inf-0.5605)': sick (3.0)
| | | | | | T4U_6 = '(0.5605-inf'
| | | | | | | T4U_9 = '(-inf-0.71575)': negative (3.0)
| | | | | | | T4U_9 = '(0.71575-inf'
| | | | | | | age_6 = '(-inf-69.1]'
| | | | | | | | TSH_1 = '(-inf-13.254875)': negative (16.73)
| | | | | | | | TSH_1 = '(13.254875-inf)': sick (6.27/1.27)
| | | | | | | | age_6 = '(69.1-inf'
| | | | | | | | on_thyroxine = f: sick (13.0/2.0)
| | | | | | | | on_thyroxine = t: negative (2.0)
| | | | | | | | T4U_measured = f: negative (3.0)
| | | | | | sick = t: sick (3.0)
| | | | | | referral_source = SVI
| | | | | | FTI_16 = '(-inf-159.2]'
| | | | | | FTI_6 = '(-inf-60.95)': negative (2.1/0.1)
| | | | | | FTI_6 = '(60.95-inf)': sick (157.32)
| | | | | | FTI_16 = '(159.2-inf'

```

```

|   |   |   |   |   TT4_9 = '(-inf-98.3)': sick (2.15)
|   |   |   |   TT4_9 = '(98.3-inf)': negative (10.44/0.44)
|   |   |   referral_source = STMW: sick (0.0)
|   |   |   referral_source = SVHD: sick (3.0)
|   T3_measured = f: negative (79.64/0.31)
T3_4 = '(1.105-inf)'
|   FTI_17 = '(-inf-169.025)': negative (3222.3/8.69)
|   FTI_17 = '(169.025-inf)'
|   |   T3_10 = '(-inf-2.6875)'
|   |   |   on_thyroxine = f
|   |   |   T3_measured = t
|   |   |   TT4_14 = '(-inf-151.8)': negative (11.57)
|   |   |   TT4_14 = '(151.8-inf)'
|   |   |   |   TT4_18 = '(-inf-194.6)': sick (20.78/4.78)
|   |   |   |   TT4_18 = '(194.6-inf)': negative (3.09)
|   |   |   T3_measured = f: negative (11.21)
|   |   |   on_thyroxine = t: negative (25.63)
|   |   T3_10 = '(2.6875-inf)': negative (86.77)

Number of Leaves : 29
Size of the tree : 54

Time taken to build model: 0.35 seconds

== Evaluation on test split ==

Time taken to test model on test split: 0.01 seconds

== Summary ==
Correctly Classified Instances      1260          98.2839 %
Incorrectly Classified Instances    22           1.7161 %
Kappa statistic                      0.8293
Mean absolute error                  0.0241
Root mean squared error              0.1245
Relative absolute error              21.0937 %
Root relative squared error         53.023 %
Total Number of Instances           1282

== Detailed Accuracy By Class ==
      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC     ROC Area  PRC Area  Class
      0.997    0.240    0.985     0.997    0.991     0.834    0.889    0.985    negative
      0.760    0.003    0.934     0.760    0.838     0.834    0.889    0.788    sick
Weighted Avg.                     0.983    0.226    0.982     0.983    0.982     0.834    0.889    0.974

== Confusion Matrix ==
      a     b  <-- classified as
1203   4 |     a = negative
     18  57 |     b = sick

```

## Leave one out fold

```

== Run information ==

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-B40-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  258
             [list of attributes omitted]
Test mode:   3772-fold cross-validation

== Classifier model (full training set) ==

J48 pruned tree
-----
T3_4 = '(-inf-1.105'
|   T3_measured = t
|   |   TT4_3 = '(-inf-34.1)': negative (42.0/1.0)
|   |   TT4_3 = '(34.1-inf)'
|   |   |   TT4_5 = '(-inf-55.5'
|   |   |   |   TSH_2 = '(-inf-26.50475'
|   |   |   |   |   FTI_8 = '(-inf-80.6'
|   |   |   |   |   |   age_4 = '(-inf-46.4)': sick (5.0)
|   |   |   |   |   |   age_4 = '(46.4-inf'
|   |   |   |   |   |   sex = F: negative (10.0/1.0)
|   |   |   |   |   |   sex = M: sick (6.0/1.0)
|   |   |   |   |   FTI_8 = '(80.6-inf)': negative (7.0
|   |   |   |   TSH_2 = '(26.50475-inf)': negative (7.0
|   |   |   TT4_5 = '(55.5-inf'

```

```

|   |   |   |   referral_source = SVHC: sick (9.0)
|   |   |   |   referral_source = other
|   |   |   |   sick = f
|   |   |   |   |   T4U_measured = t
|   |   |   |   |   |   T4U_6 = '(-inf-0.5605]': sick (3.0)
|   |   |   |   |   |   T4U_6 = '(0.5605-inf)'
|   |   |   |   |   |   T4U_9 = '(-inf-0.71575]': negative (3.0)
|   |   |   |   |   |   T4U_9 = '(0.71575-inf)'
|   |   |   |   |   |   age_6 = '(-inf-69.1]'
|   |   |   |   |   |   |   TSH_1 = '(-inf-13.254875]': negative (16.73)
|   |   |   |   |   |   |   TSH_1 = '(13.254875-inf)': sick (6.27/1.27)
|   |   |   |   |   |   age_6 = '(69.1-inf)'
|   |   |   |   |   |   |   on_thyroxine = f: sick (13.0/2.0)
|   |   |   |   |   |   |   on_thyroxine = t: negative (2.0)
|   |   |   |   |   |   T4U_measured = f: negative (3.0)
|   |   |   |   |   sick = t: sick (3.0)
|   |   |   |   |   referral_source = SVI
|   |   |   |   |   FTI_16 = '(-inf-159.2]'
|   |   |   |   |   |   FTI_6 = '(-inf-60.95]': negative (2.1/0.1)
|   |   |   |   |   |   FTI_6 = '(60.95-inf)': sick (157.32)
|   |   |   |   |   |   FTI_16 = '(159.2-inf)'
|   |   |   |   |   |   TT4_9 = '(-inf-98.3]': sick (2.15)
|   |   |   |   |   |   TT4_9 = '(98.3-inf)': negative (10.44/0.44)
|   |   |   |   |   |   referral_source = STMW: sick (0.0)
|   |   |   |   |   |   referral_source = SVHD: sick (3.0)
|   |   |   |   |   T3_measured = f: negative (79.64/0.31)
T3_4 = '(1.105-inf'
|   FTI_17 = '(-inf-169.025]': negative (3222.3/8.69)
|   FTI_17 = '(169.025-inf'
|   |   T3_10 = '(-inf-2.6875]'
|   |   |   on_thyroxine = f
|   |   |   |   T3_measured = t
|   |   |   |   |   TT4_14 = '(-inf-151.8]': negative (11.57)
|   |   |   |   |   TT4_14 = '(151.8-inf'
|   |   |   |   |   |   TT4_18 = '(-inf-194.6]': sick (20.78/4.78)
|   |   |   |   |   |   TT4_18 = '(194.6-inf)': negative (3.09)
|   |   |   |   |   T3_measured = f: negative (11.21)
|   |   |   |   |   on_thyroxine = t: negative (25.63)
|   |   |   |   T3_10 = '(2.6875-inf)': negative (86.77)

```

Number of Leaves : 29

Size of the tree : 54

Time taken to build model: 0.36 seconds

==== Stratified cross-validation ===

==== Summary ===

Correctly Classified Instances	3730	98.8865 %
Incorrectly Classified Instances	42	1.1135 %
Kappa statistic	0.9024	
Mean absolute error	0.0157	
Root mean squared error	0.0993	
Relative absolute error	13.5845 %	
Root relative squared error	41.416 %	
Total Number of Instances	3772	

==== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
0.995	0.100	0.994	0.995	0.994	0.902	0.945	0.989		negative
0.900	0.005	0.916	0.900	0.908	0.902	0.945	0.922		sick
Weighted Avg.	0.989	0.094	0.989	0.989	0.989	0.902	0.945	0.985	

==== Confusion Matrix ===

a	b	<-- classified as
3522	19	a = negative
23	208	b = sick

#### 4.1.5 Conclusion

Discretizing with equal width binning and 40 bins shows impressive results winning in 5 out 7 tests which shows that C4.5 can benefit from this preprocess step. other bins didn't show any improvements in performance.

C4.5								
Preprocessing method		Cross Validation			Percentage Split			Leave One Out Fold
		5 folds	10 folds	15 folds	50.00%	60.00%	66.00%	
With Original Dataset		98.4358%	98.8070%	98.6479%	98.3563%	98.6083%	98.0499%	98.5154%
Equal width binning	40 bins	98.4358%	98.6479%	98.9130%	98.2503%	98.6746%	98.2839%	98.8865%
	10 bins	98.1707%	98.1972%	98.1972%	98.2503%	98.1445%	97.4259%	98.0382%
	5 bins	93.8759%	93.8759%	93.8759%	93.7964%	94.1021%	94.1498%	93.8759%
	2 bins	93.8759%	93.8759%	93.8759%	93.7964%	94.1021%	94.1498%	93.8759%

## 4.2 Non supervised discretization and Equal frequency binning

### 4.2.1 2 bins

#### Cross Validation 5 folds

```
==== Run information ====
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B2-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  30
              age_1
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              T131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              T3_measured
              T3_1
              TT4_measured
              TT4_1
              T4U_measured
              T4U_1
              FTI_measured
              FTI_1
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   5-fold cross-validation

==== Classifier model (full training set) ====
J48 pruned tree
-----
: negative (3772.0/231.0)

Number of Leaves : 1
Size of the tree : 1

Time taken to build model: 0.07 seconds

==== Stratified cross-validation ====
==== Summary ====
Correctly Classified Instances      3541          93.8759 %
Incorrectly Classified Instances   231           6.1241 %
Kappa statistic                   0
Mean absolute error               0.115
Root mean squared error           0.2398
Relative absolute error           99.7787 %
Root relative squared error      99.9999 %
Total Number of Instances         3772

==== Detailed Accuracy By Class ====

```

```

TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
1.000    1.000    0.939     1.000    0.968     ?       0.498     0.939    negative
0.000    0.000    ?         0.000    ?         ?       0.498     0.061    sick
Weighted Avg. 0.939    0.939    ?         0.939    ?         ?       0.498     0.885

==== Confusion Matrix ====

a      b  <-- classified as
3541  0 |  a = negative
 231  0 |  b = sick

```

## Cross Validation 10 folds

```

==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B2-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  30
              age_1
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              i131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              T3_measured
              T3_1
              TT4_measured
              TT4_1
              T4U_measured
              T4U_1
              FTI_measured
              FTI_1
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   10-fold cross-validation

==== Classifier model (full training set) ===

J48 pruned tree
-----
: negative (3772.0/231.0)

Number of Leaves : 1

Size of the tree : 1

Time taken to build model: 0.09 seconds

==== Stratified cross-validation ===
==== Summary ===

Correctly Classified Instances      3541          93.8759 %
Incorrectly Classified Instances   231           6.1241 %
Kappa statistic                   0
Mean absolute error               0.115
Root mean squared error          0.2398
Relative absolute error           99.8032 %
Root relative squared error      99.9999 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ===

TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
1.000    1.000    0.939     1.000    0.968     ?       0.498     0.938    negative
0.000    0.000    ?         0.000    ?         ?       0.498     0.061    sick
Weighted Avg. 0.939    0.939    ?         0.939    ?         ?       0.498     0.885

```

```
==== Confusion Matrix ====

```

a	b	<- classified as
3541	0	a = negative
231	0	b = sick

## Cross Validation 15 folds

```
==== Run information ====

```

Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2  
 Relation: sick-weka.filters.unsupervised.attribute.Discretize-D-F-B2-M-1.0-Rfirst-last-precision6  
 Instances: 3772

Attributes: 30  
 age\_1  
 sex  
 on\_thyroxine  
 query\_on\_thyroxine  
 on\_antithyroid\_medication  
 sick  
 pregnant  
 thyroid\_surgery  
 i131\_treatment  
 query\_hypothyroid  
 query\_hyperthyroid  
 lithium  
 goitre  
 tumor  
 hypopituitary  
 psych  
 TSH\_measured  
 TSH\_1  
 T3\_measured  
 T3\_1  
 TT4\_measured  
 TT4\_1  
 T4U\_measured  
 T4U\_1  
 FTI\_measured  
 FTI\_1  
 TBG\_measured  
 TBG  
 referral\_source  
 Class

Test mode: 15-fold cross-validation

```
==== Classifier model (full training set) ====

```

J48 pruned tree

-----

: negative (3772.0/231.0)

Number of Leaves : 1

Size of the tree : 1

Time taken to build model: 0.06 seconds

```
==== Stratified cross-validation ====

```

```
==== Summary ====

```

	Correctly Classified Instances	3541	93.8759 %
Incorrectly Classified Instances	231	6.1241 %	
Kappa statistic	0		
Mean absolute error	0.115		
Root mean squared error	0.2398		
Relative absolute error	99.8102 %		
Root relative squared error	99.9999 %		
Total Number of Instances	3772		

```
==== Detailed Accuracy By Class ====

```

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
1.000	1.000	0.939	1.000	0.968	?	?	0.492	0.938	negative
0.000	0.000	?	0.000	?	?	?	0.492	0.060	sick
Weighted Avg.	0.939	0.939	?	0.939	?	?	0.492	0.884	

```
==== Confusion Matrix ====

```

a	b	<- classified as
3541	0	a = negative
231	0	b = sick

## Percentage Split 50%

```
==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B2-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  30
              age_1
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              T3_measured
              T3_1
              TT4_measured
              TT4_1
              T4U_measured
              T4U_1
              FTI_measured
              FTI_1
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 50.0% train, remainder test

==== Classifier model (full training set) ===

J48 pruned tree
-----
: negative (3772.0/231.0)

Number of Leaves : 1
Size of the tree : 1

Time taken to build model: 0.05 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

==== Summary ===

  Correctly Classified Instances      1769           93.7964 %
  Incorrectly Classified Instances   117            6.2036 %
  Kappa statistic                   0
  Mean absolute error               0.115
  Root mean squared error          0.2412
  Relative absolute error           99.6465 %
  Root relative squared error     100.0011 %
  Total Number of Instances        1886

==== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC    ROC Area  PRC Area  Class
      1.000    1.000    0.938     1.000    0.968     ?      0.500    0.938    negative
      0.000    0.000    ?         0.000    ?         ?      0.500    0.062    sick
  Weighted Avg.    0.938    0.938    ?         0.938    ?         ?      0.500    0.884

==== Confusion Matrix ====

      a      b  <-- classified as
1769    0 |    a = negative
      117   0 |    b = sick
```

## Percentage Split 60%

```
==== Run information ===
```

```

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B2-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes: 30
            age_1
            sex
            on_thyroxine
            query_on_thyroxine
            on_antithyroid_medication
            sick
            pregnant
            thyroid_surgery
            t131_treatment
            query_hypothyroid
            query_hyperthyroid
            lithium
            goitre
            tumor
            hypopituitary
            psych
            TSH_measured
            TSH_1
            T3_measured
            T3_1
            TT4_measured
            TT4_1
            T4U_measured
            T4U_1
            FTI_measured
            FTI_1
            TBG_measured
            TBG
            referral_source
            Class
Test mode:  split 60.0% train, remainder test

==== Classifier model (full training set) ===

J48 pruned tree
-----
: negative (3772.0/231.0)

Number of Leaves : 1

Size of the tree : 1

Time taken to build model: 0.05 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

==== Summary ===

    Correctly Classified Instances      1420           94.1021 %
    Incorrectly Classified Instances     89            5.8979 %
    Kappa statistic                      0
    Mean absolute error                  0.1143
    Root mean squared error              0.2356
    Relative absolute error              99.703 %
    Root relative squared error         99.9972 %
    Total Number of Instances           1509

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      1.000     1.000     0.941      1.000     0.970      ?       0.500      0.941   negative
      0.000     0.000     ?          0.000     ?          ?       0.500      0.059   sick
Weighted Avg.   0.941     0.941     ?          0.941     ?          ?       0.500      0.889

==== Confusion Matrix ===

  a   b   <-- classified as
1420  0 |   a = negative
  89  0 |   b = sick

```

## Percentage Split 66%

```

==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B2-M-1.0-Rfirst-last-precision6

```

```

Instances:      3772
Attributes:    30
               age_1
               sex
               on_thyroxine
               query_on_thyroxine
               on_antithyroid_medication
               sick
               pregnant
               thyroid_surgery
               I131_treatment
               query_hypothyroid
               query_hyperthyroid
               lithium
               goitre
               tumor
               hypopituitary
               psych
               TSH_measured
               TSH_1
               T3_measured
               T3_1
               TT4_measured
               TT4_1
               T4U_measured
               T4U_1
               FTI_measured
               FTI_1
               TBG_measured
               TBG
               referral_source
               Class
Test mode:     split 66.0% train, remainder test

==== Classifier model (full training set) ====

J48 pruned tree
-----
: negative (3772.0/231.0)

Number of Leaves : 1

Size of the tree : 1

Time taken to build model: 0.07 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

==== Summary ===

Correctly Classified Instances      1207          94.1498 %
Incorrectly Classified Instances   75           5.8502 %
Kappa statistic                   0
Mean absolute error              0.1138
Root mean squared error          0.2347
Relative absolute error          99.7284 %
Root relative squared error     99.9972 %
Total Number of Instances        1282

==== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
      1.000    1.000    0.941     1.000    0.970     ?       0.500    0.941    negative
      0.000    0.000    ?         0.000    ?         ?       0.500    0.059    sick
Weighted Avg.    0.941    0.941    ?         0.941    ?         ?       0.500    0.890

==== Confusion Matrix ===

      a      b  <-- classified as
1207  0 |  a = negative
      75  0 |  b = sick

```

## Leave one out fold

```

==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B2-M-1.0-Rfirst-last-precision6
Instances:    3772
Attributes:  30
               age_1

```

```

sex
on_thyroxine
query_on_thyroxine
on_antithyroid_medication
sick
pregnant
thyroid_surgery
I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH_1
T3_measured
T3_1
TT4_measured
TT4_1
T4U_measured
T4U_1
FTI_measured
FTI_1
TBG_measured
TBG
referral_source
Class
Test mode: 3772-fold cross-validation

==== Classifier model (full training set) ====
J48 pruned tree
-----
: negative (3772.0/231.0)

Number of Leaves : 1
Size of the tree : 1

Time taken to build model: 0.05 seconds

==== Stratified cross-validation ====
==== Summary ===

Correctly Classified Instances      3541          93.8759 %
Incorrectly Classified Instances   231           6.1241 %
Kappa statistic                   0
Mean absolute error               0.115
Root mean squared error          0.2398
Relative absolute error           99.8229 %
Root relative squared error      100           %
Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC    ROC Area  PRC Area  Class
      1.000    1.000    0.939     1.000    0.968     ?      0.000    0.939    negative
      0.000    0.000    ?         0.000    ?         ?      0.000    0.061    sick
Weighted Avg.    0.939    0.939    ?         0.939    ?         ?      0.000    0.885

==== Confusion Matrix ====

  a     b  <-- classified as
3541  0 |  a = negative
  231  0 |  b = sick

```

## 4.2.2 5 bins

### Cross Validation 5 folds

```

==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B5-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  48
              age_1
              age_2
              age_3
              age_4

```

```

sex
on_thyroxine
query_on_thyroxine
on_antithyroid_medication
sick
pregnant
thyroid_surgery
I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH_1
TSH_2
TSH_3
TSH_4
T3_measured
T3_1
T3_2
T3_3
T3_4
TT4_measured
TT4_1
TT4_2
TT4_3
TT4_4
T4U_measured
T4U_1
T4U_2
T4U_3
T4U_4
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
TBG_measured
TBG
referral_source
Class
Test mode: 5-fold cross-validation

==== Classifier model (full training set) ====

J48 pruned tree
-----
T3_1 = '(-inf-1.47]'
| T3_measured = t
| | referral_source = SVHC
| | | FTI_3 = '(-inf-113.5]'
| | | | age_3 = '(-inf-59.5]': negative (11.0/2.0)
| | | | age_3 = '(59.5-inf)': sick (8.0/1.0)
| | | | FTI_3 = '(113.5-inf)': negative (12.0)
| | referral_source = other
| | | age_4 = '(-inf-70.5]': negative (136.0/8.0)
| | | age_4 = '(70.5-inf]'
| | | | TT4_1 = '(-inf-83.5]'
| | | | | TSH_1 = '(-inf-0.295]': negative (4.24/1.24)
| | | | | TSH_1 = '(0.295-inf)'
| | | | | | TSH_3 = '(-inf-1.75]': sick (4.24)
| | | | | | TSH_3 = '(1.75-inf)'
| | | | | | on_thyroxine = f
| | | | | | | T4U_1 = '(-inf-0.855]': negative (3.0/1.0)
| | | | | | | T4U_1 = '(0.855-inf)': sick (4.53/1.0)
| | | | | | on_thyroxine = t: negative (2.0)
| | | | | TT4_1 = '(83.5-inf)': negative (25.0/4.0)
| | referral_source = SVI
| | | tumor = f
| | | | T4U_1 = '(-inf-0.855]'
| | | | | age_1 = '(-inf-32.5]'
| | | | | | sick = f: negative (9.94/0.48)
| | | | | | sick = t: sick (2.0)
| | | | | | age_1 = '(32.5-inf)'
| | | | | | TSH_1 = '(-inf-0.295]'
| | | | | | | age_4 = '(-inf-70.5]': sick (18.48/3.0)
| | | | | | | age_4 = '(70.5-inf)': negative (10.66/3.17)
| | | | | | | TSH_1 = '(0.295-inf)'
| | | | | | | FTI_1 = '(-inf-89.5]'
| | | | | | | | query_hypothyroid = f
| | | | | | | | | TSH_2 = '(-inf-0.995]': negative (2.12/0.12)
| | | | | | | | | TSH_2 = '(0.995-inf)'
| | | | | | | | | TSH_3 = '(-inf-1.75]': sick (2.12)

```

```

| | | | | | | | | TSH_3 = '(1.75-inf)'
| | | | | | | | | age_4 = '(-inf-70.5)': negative (18.82/7.24)
| | | | | | | | | age_4 = '(70.5-inf)': sick (7.12/2.0)
| | | | | | query_hypothyroid = t: negative (4.12/0.12)
| | | | | | FTI_1 = '(89.5-inf)': sick (106.87/35.83)
| | | | | | T4U_1 = '(0.855-inf)'
| | | | | | query_hypothyroid = f
| | | | | | TT4_4 = '(-inf-131.5]'
| | | | | | TSH_1 = '(-inf-0.295]'
| | | | | | TT4_3 = '(-inf-111.5]'
| | | | | | age_2 = '(-inf-47.5)': sick (6.03/1.03)
| | | | | | age_2 = '(47.5-inf)': negative (8.52/2.0)
| | | | | | TT4_3 = '(111.5-inf)': sick (5.52/1.0)
| | | | | | TSH_1 = '(0.295-inf)': negative (139.15/40.06)
| | | | | | TT4_4 = '(131.5-inf)': negative (18.03/4.03)
| | | | | | query_hypothyroid = t
| | | | | | T4U_4 = '(-inf-1.115]'
| | | | | | FTI_3 = '(-inf-113.5)': sick (9.29/1.0)
| | | | | | FTI_3 = '(113.5-inf)'
| | | | | | TSH_4 = '(-inf-3.15)': negative (3.0)
| | | | | | TSH_4 = '(3.15-inf)': sick (2.16)
| | | | | | T4U_4 = '(1.115-inf)': negative (2.06/0.06)
| | | | | | tumor = t: sick (2.0)
| | | | | | referral_source = STMW: negative (0.0)
| | | | | | referral_source = SVHD: sick (4.0/1.0)
| | T3_measured = f: negative (151.6/0.59)
T3_1 = '(1.47-inf)': negative (3028.4/21.41)

```

Number of Leaves : 34

Size of the tree : 64

Time taken to build model: 0.15 seconds

==== Stratified cross-validation ====  
==== Summary ===

Correctly Classified Instances	3572	94.6978 %
Incorrectly Classified Instances	200	5.3022 %
Kappa statistic	0.4938	
Mean absolute error	0.069	
Root mean squared error	0.1989	
Relative absolute error	59.8794 %	
Root relative squared error	82.9543 %	
Total Number of Instances	3772	

==== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
0.978	0.528	0.966	0.978	0.972	0.497	0.871	0.980		negative
0.472	0.022	0.583	0.472	0.522	0.497	0.871	0.453		sick
Weighted Avg.	0.947	0.497	0.943	0.947	0.944	0.497	0.871	0.947	

==== Confusion Matrix ===

a	b	<-- classified as
3463	78	a = negative
122	109	b = sick

## Cross Validation 10 folds

==== Run information ===

Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2  
 Relation: sick-weka.filters.unsupervised.attribute.Discretize-D-F-B5-M-1.0-Rfirst-last-precision6  
 Instances: 3772  
 Attributes: 48  
 age\_1  
 age\_2  
 age\_3  
 age\_4  
 sex  
 on\_thyroxine  
 query\_on\_thyroxine  
 on\_antithyroid\_medication  
 sick  
 pregnant  
 thyroid\_surgery  
 i131\_treatment  
 query\_hypothyroid  
 query\_hyperthyroid  
 lithium  
 goitre

```

tumor
hypopituitary
psych
TSH_measured
TSH_1
TSH_2
TSH_3
TSH_4
T3_measured
T3_1
T3_2
T3_3
T3_4
TT4_measured
TT4_1
TT4_2
TT4_3
TT4_4
T4U_measured
T4U_1
T4U_2
T4U_3
T4U_4
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
TBG_measured
TBG
referral_source
Class
Test mode: 10-fold cross-validation

==== Classifier model (full training set) ====

J48 pruned tree
-----
T3_1 = '(-inf-1.47]'
| T3_measured = t
| | referral_source = SVHC
| | | FTI_3 = '(-inf-113.5]'
| | | | age_3 = '(-inf-59.5]': negative (11.0/2.0)
| | | | age_3 = '(59.5-inf)': sick (8.0/1.0)
| | | | FTI_3 = '(113.5-inf)': negative (12.0)
| | | referral_source = other
| | | | age_4 = '(-inf-70.5)': negative (136.0/8.0)
| | | | age_4 = '(70.5-inf)'
| | | | | TT4_1 = '(-inf-83.5]'
| | | | | | TSH_1 = '(-inf-0.295]': negative (4.24/1.24)
| | | | | | TSH_1 = '(0.295-inf)'
| | | | | | | TSH_3 = '(-inf-1.75)': sick (4.24)
| | | | | | | TSH_3 = '(1.75-inf)'
| | | | | | | on_thyroxine = f
| | | | | | | | T4U_1 = '(-inf-0.855)': negative (3.0/1.0)
| | | | | | | | T4U_1 = '(0.855-inf)': sick (4.53/1.0)
| | | | | | | | on_thyroxine = t: negative (2.0)
| | | | | | | TT4_1 = '(83.5-inf)': negative (25.0/4.0)
| | | | | | | referral_source = SVI
| | | | | | | tumor = f
| | | | | | | T4U_1 = '(-inf-0.855)'
| | | | | | | age_1 = '(-inf-32.5]'
| | | | | | | | sick = f: negative (9.94/0.48)
| | | | | | | | sick = t: sick (2.0)
| | | | | | | | age_1 = '(32.5-inf)'
| | | | | | | | TSH_1 = '(-inf-0.295]'
| | | | | | | | | age_4 = '(-inf-70.5)': sick (18.48/3.0)
| | | | | | | | | age_4 = '(70.5-inf)': negative (10.66/3.17)
| | | | | | | | | TSH_1 = '(0.295-inf)'
| | | | | | | | | FTI_1 = '(-inf-89.5]'
| | | | | | | | | | query_hypothyroid = f
| | | | | | | | | | TSH_2 = '(-inf-0.995)': negative (2.12/0.12)
| | | | | | | | | | TSH_2 = '(0.995-inf)'
| | | | | | | | | | | TSH_3 = '(-inf-1.75)': sick (2.12)
| | | | | | | | | | | TSH_3 = '(1.75-inf)'
| | | | | | | | | | | | age_4 = '(-inf-70.5)': negative (18.82/7.24)
| | | | | | | | | | | | age_4 = '(70.5-inf)': sick (7.12/2.0)
| | | | | | | | | | | | query_hypothyroid = t: negative (4.12/0.12)
| | | | | | | | | | | | FTI_1 = '(89.5-inf)': sick (106.87/35.83)
| | | | | | | | | | | | T4U_1 = '(0.855-inf)'
| | | | | | | | | | | | | query_hypothyroid = f
| | | | | | | | | | | | | TT4_4 = '(-inf-131.5]'
| | | | | | | | | | | | | TSH_1 = '(-inf-0.295]'
| | | | | | | | | | | | | TT4_3 = '(-inf-111.5]'
| | | | | | | | | | | | | | age_2 = '(-inf-47.5)': sick (6.03/1.03)
| | | | | | | | | | | | | | age_2 = '(47.5-inf)': negative (8.52/2.0)

```

```

|   |   |   |   |   |   |   TT4_3 = '(111.5-inf)': sick (5.52/1.0)
|   |   |   |   |   |   TSH_1 = '(0.295-inf)': negative (139.15/40.06)
|   |   |   |   |   TT4_4 = '(131.5-inf)': negative (18.03/4.03)
|   |   |   |   query_hypothyroid = t
|   |   |   |   T4U_4 = '(-inf-1.115]'
|   |   |   |   FTI_3 = '(-inf-113.5]': sick (9.29/1.0)
|   |   |   |   FTI_3 = '(113.5-inf)'
|   |   |   |   |   TSH_4 = '(-inf-3.15]': negative (3.0)
|   |   |   |   |   TSH_4 = '(3.15-inf)': sick (2.16)
|   |   |   |   T4U_4 = '(1.115-inf)': negative (2.06/0.06)
|   |   |   tumor = t: sick (2.0)
|   |   referral_source = STMW: negative (0.0)
|   |   referral_source = SVHD: sick (4.0/1.0)
|   T3_measured = f: negative (151.6/0.59)
T3_1 = '(1.47-inf)': negative (3028.4/21.41)

Number of Leaves : 34
Size of the tree : 64

Time taken to build model: 0.08 seconds

== Stratified cross-validation ==
== Summary ==

Correctly Classified Instances      3571          94.6713 %
Incorrectly Classified Instances    201           5.3287 %
Kappa statistic                      0.4946
Mean absolute error                  0.0685
Root mean squared error              0.1989
Relative absolute error              59.4491 %
Root relative squared error         82.9565 %
Total Number of Instances           3772

== Detailed Accuracy By Class ==

      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
0.977     0.524     0.966     0.977     0.972     0.497     0.881     0.981  negative
0.476     0.023     0.579     0.476     0.523     0.497     0.881     0.466   sick
Weighted Avg.  0.947     0.493     0.943     0.947     0.944     0.497     0.881     0.950

== Confusion Matrix ==

      a     b  <-- classified as
3461   80 |     a = negative
121   110 |     b = sick

```

## Cross Validation 15 folds

```

== Run information ==

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B5-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  48
             age_1
             age_2
             age_3
             age_4
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             t131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary
             psych
             TSH_measured
             TSH_1
             TSH_2
             TSH_3
             TSH_4
             T3_measured
             T3_1
             T3_2
             T3_3

```

```

T3_4
TT4_measured
TT4_1
TT4_2
TT4_3
TT4_4
T4U_measured
T4U_1
T4U_2
T4U_3
T4U_4
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
TBG_measured
TBG
referral_source
Class
Test mode: 15-fold cross-validation

==== Classifier model (full training set) ====

J48 pruned tree
-----

T3_1 = '(-inf-1.47)'
| T3_measured = t
|   | referral_source = SVHC
|   |   | FTI_3 = '(-inf-113.5)'
|   |   |   | age_3 = '(-inf-59.5)': negative (11.0/2.0)
|   |   |   | age_3 = '(59.5-inf)': sick (8.0/1.0)
|   |   |   | FTI_3 = '(113.5-inf)': negative (12.0)
|   | referral_source = other
|   |   | age_4 = '(-inf-70.5)': negative (136.0/8.0)
|   |   | age_4 = '(70.5-inf)'
|   |   |   | TT4_1 = '(-inf-83.5)'
|   |   |   |   | TSH_1 = '(-inf-0.295)': negative (4.24/1.24)
|   |   |   |   | TSH_1 = '(0.295-inf)'
|   |   |   |   | TSH_3 = '(-inf-1.75)': sick (4.24)
|   |   |   |   | TSH_3 = '(1.75-inf)'
|   |   |   |   |   | on_thyroxine = f
|   |   |   |   |   |   | T4U_1 = '(-inf-0.855)': negative (3.0/1.0)
|   |   |   |   |   |   | T4U_1 = '(0.855-inf)': sick (4.53/1.0)
|   |   |   |   |   |   |   | on_thyroxine = t: negative (2.0)
|   |   |   |   |   |   |   | TT4_1 = '(83.5-inf)': negative (25.0/4.0)
|   | referral_source = SVI
|   | tumor = f
|   |   | T4U_1 = '(-inf-0.855)'
|   |   | age_1 = '(-inf-32.5)'
|   |   |   | sick = f: negative (9.94/0.48)
|   |   |   | sick = t: sick (2.0)
|   |   |   | age_1 = '(32.5-inf)'
|   |   |   |   | TSH_1 = '(-inf-0.295)'
|   |   |   |   | age_4 = '(-inf-70.5)': sick (18.48/3.0)
|   |   |   |   | age_4 = '(70.5-inf)': negative (10.66/3.17)
|   |   |   |   | TSH_1 = '(0.295-inf)'
|   |   |   |   |   | FTI_1 = '(-inf-89.5)'
|   |   |   |   |   |   | query_hypothyroid = f
|   |   |   |   |   |   |   | TSH_2 = '(-inf-0.995)': negative (2.12/0.12)
|   |   |   |   |   |   |   | TSH_2 = '(0.995-inf)'
|   |   |   |   |   |   |   |   | TSH_3 = '(-inf-1.75)': sick (2.12)
|   |   |   |   |   |   |   |   | TSH_3 = '(1.75-inf)'
|   |   |   |   |   |   |   |   | age_4 = '(-inf-70.5)': negative (18.82/7.24)
|   |   |   |   |   |   |   |   | age_4 = '(70.5-inf)': sick (7.12/2.0)
|   |   |   |   |   |   |   |   | query_hypothyroid = t: negative (4.12/0.12)
|   |   |   |   |   |   |   |   |   | FTI_1 = '(89.5-inf)': sick (106.87/35.83)
|   |   |   |   |   |   |   |   | T4U_1 = '(0.855-inf)'
|   |   |   |   |   |   |   |   | query_hypothyroid = f
|   |   |   |   |   |   |   |   | TT4_4 = '(-inf-131.5)'
|   |   |   |   |   |   |   |   |   | TSH_1 = '(-inf-0.295)'
|   |   |   |   |   |   |   |   |   | TT4_3 = '(-inf-111.5)'
|   |   |   |   |   |   |   |   |   |   | age_2 = '(-inf-47.5)': sick (6.03/1.03)
|   |   |   |   |   |   |   |   |   |   | age_2 = '(47.5-inf)': negative (8.52/2.0)
|   |   |   |   |   |   |   |   |   |   | TT4_3 = '(111.5-inf)': sick (5.52/1.0)
|   |   |   |   |   |   |   |   |   |   | TSH_1 = '(0.295-inf)': negative (139.15/40.06)
|   |   |   |   |   |   |   |   |   |   | TT4_4 = '(131.5-inf)': negative (18.03/4.03)
|   |   |   |   |   |   |   |   |   |   | query_hypothyroid = t
|   |   |   |   |   |   |   |   |   |   |   | T4U_4 = '(-inf-1.115)'
|   |   |   |   |   |   |   |   |   |   |   |   | FTI_3 = '(-inf-113.5)': sick (9.29/1.0)
|   |   |   |   |   |   |   |   |   |   |   |   | FTI_3 = '(113.5-inf)'
|   |   |   |   |   |   |   |   |   |   |   |   |   | TSH_4 = '(-inf-3.15)': negative (3.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   | TSH_4 = '(3.15-inf)': sick (2.16)
|   |   |   |   |   |   |   |   |   |   |   |   |   | T4U_4 = '(1.115-inf)': negative (2.06/0.06)
|   | tumor = t: sick (2.0)
|   | referral_source = STMW: negative (0.0)

```

```

|   |   referral_source = SVHD: sick (4.0/1.0)
|   T3_measured = f: negative (151.6/0.59)
T3_1 = '(1.47-inf)': negative (3028.4/21.41)

Number of Leaves : 34

Size of the tree : 64

Time taken to build model: 0.09 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3572          94.6978 %
Incorrectly Classified Instances   200           5.3022 %
Kappa statistic                   0.5048
Mean absolute error               0.0692
Root mean squared error          0.1985
Relative absolute error          60.0517 %
Root relative squared error     82.7705 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC   ROC Area   PRC Area   Class
0.977    0.506    0.967    0.977    0.972    0.507    0.895    0.986    negative
0.494    0.023    0.579    0.494    0.533    0.507    0.895    0.461    sick
Weighted Avg.   0.947    0.477    0.943    0.947    0.945    0.507    0.895    0.954

==== Confusion Matrix ====

      a      b  <-- classified as
3458   83 |      a = negative
117    114 |      b = sick

```

## Percentage Split 50%

```

==== Run information ====

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B5-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  48
              age_1
              age_2
              age_3
              age_4
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              TSH_2
              TSH_3
              TSH_4
              T3_measured
              T3_1
              T3_2
              T3_3
              T3_4
              TT4_measured
              TT4_1
              TT4_2
              TT4_3
              TT4_4
              T4U_measured
              T4U_1
              T4U_2
              T4U_3
              T4U_4
              FTI_measured

```

```

FTI_1
FTI_2
FTI_3
FTI_4
TBG_measured
TBG
referral_source
Class
Test mode: split 50.0% train, remainder test

==== Classifier model (full training set) ===

J48 pruned tree
-----

T3_1 = '(-inf-1.47]'
| T3_measured = t
| | referral_source = SVHC
| | | FTI_3 = '(-inf-113.5]'
| | | | age_3 = '(-inf-59.5]': negative (11.0/2.0)
| | | | age_3 = '(59.5-inf)': sick (8.0/1.0)
| | | | FTI_3 = '(113.5-inf)': negative (12.0)
| | | | referral_source = other
| | | | age_4 = '(-inf-70.5)': negative (136.0/8.0)
| | | | age_4 = '(70.5-inf)'
| | | | | TT4_1 = '(-inf-83.5]'
| | | | | TSH_1 = '(-inf-0.295)': negative (4.24/1.24)
| | | | | TSH_1 = '(0.295-inf)'
| | | | | TSH_3 = '(-inf-1.75)': sick (4.24)
| | | | | TSH_3 = '(1.75-inf)'
| | | | | on_thyroxine = f
| | | | | | T4U_1 = '(-inf-0.855)': negative (3.0/1.0)
| | | | | | T4U_1 = '(0.855-inf)': sick (4.53/1.0)
| | | | | | on_thyroxine = t: negative (2.0)
| | | | | | TT4_1 = '(83.5-inf)': negative (25.0/4.0)
| | | | | referral_source = SVI
| | | | | tumor = f
| | | | | T4U_1 = '(-inf-0.855'
| | | | | | age_1 = '(-inf-32.5]'
| | | | | | | sick = f: negative (9.94/0.48)
| | | | | | | sick = t: sick (2.0)
| | | | | | | age_1 = '(32.5-inf)'
| | | | | | | TSH_1 = '(-inf-0.295]'
| | | | | | | | age_4 = '(-inf-70.5)': sick (18.48/3.0)
| | | | | | | | age_4 = '(70.5-inf)': negative (10.66/3.17)
| | | | | | | TSH_1 = '(0.295-inf)'
| | | | | | | | FTI_1 = '(-inf-89.5]'
| | | | | | | | | query_hypothyroid = f
| | | | | | | | | TSH_2 = '(-inf-0.995)': negative (2.12/0.12)
| | | | | | | | | TSH_2 = '(0.995-inf)'
| | | | | | | | | | TSH_3 = '(-inf-1.75)': sick (2.12)
| | | | | | | | | | TSH_3 = '(1.75-inf)'
| | | | | | | | | | age_4 = '(-inf-70.5)': negative (18.82/7.24)
| | | | | | | | | | age_4 = '(70.5-inf)': sick (7.12/2.0)
| | | | | | | | | | query_hypothyroid = t: negative (4.12/0.12)
| | | | | | | | | | FTI_1 = '(89.5-inf)': sick (106.87/35.83)
| | | | | | | | T4U_1 = '(0.855-inf)'
| | | | | | | | query_hypothyroid = f
| | | | | | | | TT4_4 = '(-inf-131.5]'
| | | | | | | | | TSH_1 = '(-inf-0.295]'
| | | | | | | | | | TT4_3 = '(-inf-111.5]'
| | | | | | | | | | | age_2 = '(-inf-47.5)': sick (6.03/1.03)
| | | | | | | | | | | age_2 = '(47.5-inf)': negative (8.52/2.0)
| | | | | | | | | | | TT4_3 = '(111.5-inf)': sick (5.52/1.0)
| | | | | | | | | | | TSH_1 = '(0.295-inf)': negative (139.15/40.06)
| | | | | | | | | | | TT4_4 = '(131.5-inf)': negative (18.03/4.03)
| | | | | | | | | | | query_hypothyroid = t
| | | | | | | | | | | T4U_4 = '(-inf-1.115]'
| | | | | | | | | | | FTI_3 = '(-inf-113.5)': sick (9.29/1.0)
| | | | | | | | | | | FTI_3 = '(113.5-inf)'
| | | | | | | | | | | TSH_4 = '(-inf-3.15)': negative (3.0)
| | | | | | | | | | | TSH_4 = '(3.15-inf)': sick (2.16)
| | | | | | | | | | | T4U_4 = '(1.115-inf)': negative (2.06/0.06)
| | | | | | | | | | | tumor = t: sick (2.0)
| | | | | | | | | | | referral_source = STMW: negative (0.0)
| | | | | | | | | | | referral_source = SVHD: sick (4.0/1.0)
| | | | | | | | | | | T3_measured = f: negative (151.6/0.59)
T3_1 = '(1.47-inf)': negative (3028.4/21.41)

```

Number of Leaves : 34

Size of the tree : 64

Time taken to build model: 0.08 seconds

==== Evaluation on test split ===

```

Time taken to test model on test split: 0 seconds

==== Summary ====
Correctly Classified Instances      1783          94.5387 %
Incorrectly Classified Instances    103           5.4613 %
Kappa statistic                   0.4442
Mean absolute error               0.0684
Root mean squared error          0.2021
Relative absolute error          59.2991 %
Root relative squared error     83.7729 %
Total Number of Instances        1886

==== Detailed Accuracy By Class ====
      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.982     0.607     0.961     0.982     0.971     0.454     0.915     0.990     negative
      0.393     0.018     0.590     0.393     0.472     0.454     0.915     0.471     sick
Weighted Avg.      0.945     0.570     0.938     0.945     0.940     0.454     0.915     0.958

==== Confusion Matrix ====
      a      b  <-- classified as
1737  32 |  a = negative
    1    46 |  b = sick

```

## Percentage Split 60%

```

==== Run information ====
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B5-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  48
             age_1
             age_2
             age_3
             age_4
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             I131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary
             psych
             TSH_measured
             TSH_1
             TSH_2
             TSH_3
             TSH_4
             T3_measured
             T3_1
             T3_2
             T3_3
             T3_4
             TT4_measured
             TT4_1
             TT4_2
             TT4_3
             TT4_4
             T4U_measured
             T4U_1
             T4U_2
             T4U_3
             T4U_4
             FTI_measured
             FTI_1
             FTI_2
             FTI_3
             FTI_4
             TBG_measured
             TBG
             referral_source
             Class
Test mode:   split 60.0% train, remainder test

```

```

==== Classifier model (full training set) ====
J48 pruned tree
-----
T3_1 = '(-inf-1.47]'
| T3_measured = t
|   referral_source = SVHC
|   | FTI_3 = '(-inf-113.5]'
|   |   | age_3 = '(-inf-59.5]': negative (11.0/2.0)
|   |   | age_3 = '(59.5-inf)': sick (8.0/1.0)
|   |   | FTI_3 = '(113.5-inf)': negative (12.0)
|   referral_source = other
|   | age_4 = '(-inf-70.5)': negative (136.0/8.0)
|   | age_4 = '(70.5-inf)'
|   |   | TT4_1 = '(-inf-83.5]'
|   |   |   TSH_1 = '(-inf-0.295)': negative (4.24/1.24)
|   |   |   TSH_1 = '(0.295-inf)'
|   |   |   | TSH_3 = '(-inf-1.75)': sick (4.24)
|   |   |   | TSH_3 = '(1.75-inf)'
|   |   |   | on_thyroxine = f
|   |   |   |   T4U_1 = '(-inf-0.855)': negative (3.0/1.0)
|   |   |   |   T4U_1 = '(0.855-inf)': sick (4.53/1.0)
|   |   |   | on_thyroxine = t: negative (2.0)
|   |   |   | TT4_1 = '(83.5-inf)': negative (25.0/4.0)
referral_source = SVI
| tumor = f
|   | T4U_1 = '(-inf-0.855]'
|   |   | age_1 = '(-inf-32.5]'
|   |   |   | sick = f: negative (9.94/0.48)
|   |   |   | sick = t: sick (2.0)
|   |   |   | age_1 = '(32.5-inf)'
|   |   |   | TSH_1 = '(-inf-0.295]'
|   |   |   |   | age_4 = '(-inf-70.5)': sick (18.48/3.0)
|   |   |   |   | age_4 = '(70.5-inf)': negative (10.66/3.17)
|   |   |   |   TSH_1 = '(0.295-inf)'
|   |   |   |   | FTI_1 = '(-inf-89.5]'
|   |   |   |   |   query_hypothyroid = f
|   |   |   |   |   | TSH_2 = '(-inf-0.995)': negative (2.12/0.12)
|   |   |   |   |   | TSH_2 = '(0.995-inf)'
|   |   |   |   |   |   | TSH_3 = '(-inf-1.75)': sick (2.12)
|   |   |   |   |   |   | TSH_3 = '(1.75-inf)'
|   |   |   |   |   |   |   | age_4 = '(-inf-70.5)': negative (18.82/7.24)
|   |   |   |   |   |   |   | age_4 = '(70.5-inf)': sick (7.12/2.0)
|   |   |   |   |   |   |   | query_hypothyroid = t: negative (4.12/0.12)
|   |   |   |   |   |   |   | FTI_1 = '(89.5-inf)': sick (106.87/35.83)
|   |   |   T4U_1 = '(0.855-inf)'
|   |   |   |   query_hypothyroid = f
|   |   |   |   | TT4_4 = '(-inf-131.5]'
|   |   |   |   |   TSH_1 = '(-inf-0.295]'
|   |   |   |   |   | TT4_3 = '(-inf-111.5]'
|   |   |   |   |   |   | age_2 = '(-inf-47.5)': sick (6.03/1.03)
|   |   |   |   |   |   | age_2 = '(47.5-inf)': negative (8.52/2.0)
|   |   |   |   |   |   | TT4_3 = '(111.5-inf)': sick (5.52/1.0)
|   |   |   |   |   |   |   TSH_1 = '(0.295-inf)': negative (139.15/40.06)
|   |   |   |   |   |   |   | TT4_4 = '(131.5-inf)': negative (18.03/4.03)
|   |   |   |   |   |   |   | query_hypothyroid = t
|   |   |   |   |   |   |   | T4U_4 = '(-inf-1.115]'
|   |   |   |   |   |   |   |   | FTI_3 = '(-inf-113.5)': sick (9.29/1.0)
|   |   |   |   |   |   |   |   | FTI_3 = '(113.5-inf)'
|   |   |   |   |   |   |   |   | TSH_4 = '(-inf-3.15)': negative (3.0)
|   |   |   |   |   |   |   |   | TSH_4 = '(3.15-inf)': sick (2.16)
|   |   |   |   |   |   |   |   | T4U_4 = '(1.115-inf)': negative (2.06/0.06)
|   |   |   |   |   |   |   | tumor = t: sick (2.0)
|   |   |   |   |   |   |   | referral_source = STMW: negative (0.0)
|   |   |   |   |   |   |   | referral_source = SVHD: sick (4.0/1.0)
|   T3_measured = f: negative (151.6/0.59)
T3_1 = '(1.47-inf)': negative (3028.4/21.41)

Number of Leaves : 34

Size of the tree : 64

Time taken to build model: 0.08 seconds

==== Evaluation on test split ====

Time taken to test model on test split: 0 seconds

==== Summary ===



|                                  | Correctly Classified Instances | 1434     | 95.0298 % |
|----------------------------------|--------------------------------|----------|-----------|
| Incorrectly Classified Instances | 75                             | 4.9702 % |           |
| Kappa statistic                  | 0.4715                         |          |           |
| Mean absolute error              | 0.0636                         |          |           |


```

```

Root mean squared error          0.1924
Relative absolute error        55.4584 %
Root relative squared error    81.6676 %
Total Number of Instances      1509

==== Detailed Accuracy By Class ====

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC     ROC Area  PRC Area  Class
      0.984    0.584    0.964    0.984    0.974    0.482    0.907    0.989    negative
      0.416    0.016    0.617    0.416    0.497    0.482    0.907    0.494    sick
Weighted Avg.      0.950    0.551    0.944    0.950    0.946    0.482    0.907    0.960

==== Confusion Matrix ====

      a      b  <-- classified as
1397   23 |      a = negative
      52   37 |      b = sick

```

## Percentage Split 66%

```

==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B5-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  48
              age_1
              age_2
              age_3
              age_4
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              TSH_2
              TSH_3
              TSH_4
              T3_measured
              T3_1
              T3_2
              T3_3
              T3_4
              TT4_measured
              TT4_1
              TT4_2
              TT4_3
              TT4_4
              T4U_measured
              T4U_1
              T4U_2
              T4U_3
              T4U_4
              FTI_measured
              FTI_1
              FTI_2
              FTI_3
              FTI_4
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 66.0% train, remainder test

==== Classifier model (full training set) ===

J48 pruned tree
-----
T3_1 = '(-inf-1.47)'
|  T3_measured = t
|  |  referral_source = SVHC

```

```

|   |   |   FTI_3 = '(-inf-113.5]'
|   |   |   age_3 = '(-inf-59.5]': negative (11.0/2.0)
|   |   |   age_3 = '(59.5-inf)': sick (8.0/1.0)
|   |   |   FTI_3 = '(113.5-inf)': negative (12.0)
|   |   referral_source = other
|   |   age_4 = '(-inf-70.5]': negative (136.0/8.0)
|   |   age_4 = '(70.5-inf)'
|   |   |   TT4_1 = '(-inf-83.5]'
|   |   |   TSH_1 = '(-inf-0.295]': negative (4.24/1.24)
|   |   |   TSH_1 = '(0.295-inf)'
|   |   |   |   TSH_3 = '(-inf-1.75]': sick (4.24)
|   |   |   |   TSH_3 = '(1.75-inf)'
|   |   |   |   on_thyroxine = f
|   |   |   |   |   T4U_1 = '(-inf-0.855]': negative (3.0/1.0)
|   |   |   |   |   T4U_1 = '(0.855-inf)': sick (4.53/1.0)
|   |   |   |   on_thyroxine = t: negative (2.0)
|   |   |   TT4_1 = '(83.5-inf)': negative (25.0/4.0)
|   |   referral_source = SVI
|   |   tumor = f
|   |   |   T4U_1 = '(-inf-0.855]'
|   |   |   age_1 = '(-inf-32.5]'
|   |   |   |   sick = f: negative (9.94/0.48)
|   |   |   |   sick = t: sick (2.0)
|   |   |   age_1 = '(32.5-inf)'
|   |   |   |   TSH_1 = '(-inf-0.295]'
|   |   |   |   age_4 = '(-inf-70.5]': sick (18.48/3.0)
|   |   |   |   age_4 = '(70.5-inf)': negative (10.66/3.17)
|   |   |   TSH_1 = '(0.295-inf'
|   |   |   |   |   FTI_1 = '(-inf-89.5]'
|   |   |   |   |   query_hypothyroid = f
|   |   |   |   |   TSH_2 = '(-inf-0.995]': negative (2.12/0.12)
|   |   |   |   |   TSH_2 = '(0.995-inf)'
|   |   |   |   |   |   TSH_3 = '(-inf-1.75]': sick (2.12)
|   |   |   |   |   |   TSH_3 = '(1.75-inf)'
|   |   |   |   |   |   age_4 = '(-inf-70.5]': negative (18.82/7.24)
|   |   |   |   |   |   age_4 = '(70.5-inf)': sick (7.12/2.0)
|   |   |   |   |   query_hypothyroid = t: negative (4.12/0.12)
|   |   |   |   |   FTI_1 = '(89.5-inf)': sick (106.87/35.83)
|   |   |   T4U_1 = '(0.855-inf'
|   |   |   |   query_hypothyroid = f
|   |   |   |   TT4_4 = '(-inf-131.5]'
|   |   |   |   TSH_1 = '(-inf-0.295]'
|   |   |   |   |   TT4_3 = '(-inf-111.5]'
|   |   |   |   |   age_2 = '(-inf-47.5]': sick (6.03/1.03)
|   |   |   |   |   age_2 = '(47.5-inf)': negative (8.52/2.0)
|   |   |   |   |   TT4_3 = '(111.5-inf)': sick (5.52/1.0)
|   |   |   |   |   TSH_1 = '(0.295-inf)': negative (139.15/40.06)
|   |   |   |   TT4_4 = '(131.5-inf)': negative (18.03/4.03)
|   |   |   |   query_hypothyroid = t
|   |   |   |   T4U_4 = '(-inf-1.115]'
|   |   |   |   |   FTI_3 = '(-inf-113.5]': sick (9.29/1.0)
|   |   |   |   |   FTI_3 = '(113.5-inf)'
|   |   |   |   |   TSH_4 = '(-inf-3.15]': negative (3.0)
|   |   |   |   |   TSH_4 = '(3.15-inf)': sick (2.16)
|   |   |   |   T4U_4 = '(1.115-inf)': negative (2.06/0.06)
|   |   |   tumor = t: sick (2.0)
|   |   referral_source = STMW: negative (0.0)
|   |   referral_source = SVHD: sick (4.0/1.0)
|   T3_measured = f: negative (151.6/0.59)
T3_1 = '(1.47-inf)': negative (3028.4/21.41)

Number of Leaves : 34

Size of the tree : 64

Time taken to build model: 0.08 seconds

== Evaluation on test split ==

Time taken to test model on test split: 0 seconds

== Summary ==

Correctly Classified Instances      1212          94.5398 %
Incorrectly Classified Instances     70           5.4602 %
Kappa statistic                      0.3705
Mean absolute error                  0.0666
Root mean squared error              0.1955
Relative absolute error              58.352  %
Root relative squared error         83.2919 %
Total Number of Instances            1282

== Detailed Accuracy By Class ==

      TP Rate   FP Rate   Precision   Recall    F-Measure   MCC      ROC Area   PRC Area   Class
0.985      0.693      0.958      0.985      0.971      0.389      0.902      0.989   negative

```

```

          0.307    0.015    0.561    0.307    0.397    0.389    0.902    0.456    sick
Weighted Avg.    0.945    0.654    0.935    0.945    0.938    0.389    0.902    0.958

==== Confusion Matrix ====
      a      b  <-- classified as
1189   18 |     a = negative
      52   23 |     b = sick

```

## Leave one out fold

```

==== Run information ====
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B5-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  48
              age_1
              age_2
              age_3
              age_4
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              t131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              TSH_2
              TSH_3
              TSH_4
              T3_measured
              T3_1
              T3_2
              T3_3
              T3_4
              TT4_measured
              TT4_1
              TT4_2
              TT4_3
              TT4_4
              T4U_measured
              T4U_1
              T4U_2
              T4U_3
              T4U_4
              FTI_measured
              FTI_1
              FTI_2
              FTI_3
              FTI_4
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   3772-fold cross-validation

==== Classifier model (full training set) ====
J48 pruned tree
-----
T3_1 = '(-inf-1.47)'
|  T3_measured = t
|  |  referral_source = SVHC
|  |  |  FTI_3 = '(-inf-113.5'
|  |  |  |  age_3 = '(-inf-59.5': negative (11.0/2.0)
|  |  |  |  age_3 = '(59.5-inf)': sick (8.0/1.0)
|  |  |  |  FTI_3 = '(113.5-inf)': negative (12.0)
|  |  |  referral_source = other
|  |  |  |  age_4 = '(-inf-70.5': negative (136.0/8.0)
|  |  |  |  age_4 = '(70.5-inf'
|  |  |  |  |  TT4_1 = '(-inf-83.5'
|  |  |  |  |  |  TSH_1 = '(-inf-0.295)': negative (4.24/1.24)

```

```

| | | | |   TSH_1 = '(0.295-inf)'
| | | | |   TSH_3 = '(-inf-1.75)': sick (4.24)
| | | | |   TSH_3 = '(1.75-inf)'
| | | | |   on_thyroxine = f
| | | | |   |   T4U_1 = '(-inf-0.855)': negative (3.0/1.0)
| | | | |   |   T4U_1 = '(0.855-inf)': sick (4.53/1.0)
| | | | |   on_thyroxine = t: negative (2.0)
| | | | TT4_1 = '(83.5-inf)': negative (25.0/4.0)
| referral_source = SVI
| tumor = f
|   T4U_1 = '(-inf-0.855'
|   age_1 = '(-inf-32.5]'
|     sick = f: negative (9.94/0.48)
|     sick = t: sick (2.0)
|   age_1 = '(32.5-inf)'
|   |   TSH_1 = '(-inf-0.295'
|   |   |   age_4 = '(-inf-70.5)': sick (18.48/3.0)
|   |   |   age_4 = '(70.5-inf)': negative (10.66/3.17)
|   |   TSH_1 = '(0.295-inf)'
|   |   |   FTI_1 = '(-inf-89.5'
|   |   |   |   query_hypothyroid = f
|   |   |   TSH_2 = '(-inf-0.995)': negative (2.12/0.12)
|   |   |   TSH_2 = '(0.995-inf'
|   |   |   |   TSH_3 = '(-inf-1.75)': sick (2.12)
|   |   |   |   TSH_3 = '(1.75-inf'
|   |   |   |   age_4 = '(-inf-70.5)': negative (18.82/7.24)
|   |   |   |   age_4 = '(70.5-inf)': sick (7.12/2.0)
|   |   |   |   query_hypothyroid = t: negative (4.12/0.12)
|   |   |   FTI_1 = '(89.5-inf)': sick (106.87/35.83)
|   |   T4U_1 = '(0.855-inf'
|   |   |   query_hypothyroid = f
|   |   |   TT4_4 = '(-inf-131.5'
|   |   |   |   TSH_1 = '(-inf-0.295'
|   |   |   |   TT4_3 = '(-inf-111.5'
|   |   |   |   |   age_2 = '(-inf-47.5)': sick (6.03/1.03)
|   |   |   |   |   age_2 = '(47.5-inf)': negative (8.52/2.0)
|   |   |   |   TT4_3 = '(111.5-inf)': sick (5.52/1.0)
|   |   |   |   TSH_1 = '(0.295-inf)': negative (139.15/40.06)
|   |   |   TT4_4 = '(131.5-inf)': negative (18.03/4.03)
|   |   |   query_hypothyroid = t
|   |   |   T4U_4 = '(-inf-1.115'
|   |   |   |   FTI_3 = '(-inf-113.5)': sick (9.29/1.0)
|   |   |   |   FTI_3 = '(113.5-inf'
|   |   |   |   TSH_4 = '(-inf-3.15)': negative (3.0)
|   |   |   |   TSH_4 = '(3.15-inf)': sick (2.16)
|   |   |   T4U_4 = '(1.115-inf)': negative (2.06/0.06)
|   |   tumor = t: sick (2.0)
|   referral_source = STMW: negative (0.0)
|   referral_source = SVHD: sick (4.0/1.0)
|   T3_measured = f: negative (151.6/0.59)
T3_1 = '(1.47-inf)': negative (3028.4/21.41)

Number of Leaves : 34
Size of the tree : 64

Time taken to build model: 0.08 seconds

==== Stratified cross-validation ====
==== Summary ===

  Correctly Classified Instances      3568          94.5917 %
  Incorrectly Classified Instances    204           5.4083 %
  Kappa statistic                   0.4927
  Mean absolute error               0.0697
  Root mean squared error          0.2014
  Relative absolute error          60.5288 %
  Root relative squared error      83.9938 %
  Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.976     0.519     0.966     0.976     0.971     0.495     0.883     0.986     negative
      0.481     0.024     0.569     0.481     0.521     0.495     0.883     0.427     sick
  Weighted Avg.   0.946     0.489     0.942     0.946     0.944     0.495     0.883     0.952

==== Confusion Matrix ====

      a     b  <-- classified as
3457   84 |   a = negative
      120  111 |   b = sick

```

### 4.2.3 10 bins

#### Cross Validation 5 folds

```
==== Run information ====  
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2  
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B10-M-1.0-Rfirst-last-precision6  
Instances:   3772  
Attributes:  78  
              age_1  
              age_2  
              age_3  
              age_4  
              age_5  
              age_6  
              age_7  
              age_8  
              age_9  
              sex  
              on_thyroxine  
              query_on_thyroxine  
              on_antithyroid_medication  
              sick  
              pregnant  
              thyroid_surgery  
              I131_treatment  
              query_hypothyroid  
              query_hyperthyroid  
              lithium  
              goitre  
              tumor  
              hypopituitary  
              psych  
              TSH_measured  
              TSH_1  
              TSH_2  
              TSH_3  
              TSH_4  
              TSH_5  
              TSH_6  
              TSH_7  
              TSH_8  
              TSH_9  
              T3_measured  
              T3_1  
              T3_2  
              T3_3  
              T3_4  
              T3_5  
              T3_6  
              T3_7  
              T3_8  
              T3_9  
              TT4_measured  
              TT4_1  
              TT4_2  
              TT4_3  
              TT4_4  
              TT4_5  
              TT4_6  
              TT4_7  
              TT4_8  
              TT4_9  
              T4U_measured  
              T4U_1  
              T4U_2  
              T4U_3  
              T4U_4  
              T4U_5  
              T4U_6  
              T4U_7  
              T4U_8  
              T4U_9  
              FTI_measured  
              FTI_1  
              FTI_2  
              FTI_3  
              FTI_4  
              FTI_5  
              FTI_6  
              FTI_7  
              FTI_8  
              FTI_9  
              TBG_measured  
              TBG
```

## Cross Validation 10 folds

```
==== Run information ====  
  
Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2
```

```
Relation: sick-weka.filters.unsupervised.attribute.Discretize-D-F-B10-M-1.0-Rfirst-last-precision6
Instances: 3772
Attributes: 78
age_1
age_2
age_3
age_4
age_5
age_6
age_7
age_8
age_9
sex
on_thyroxine
query_on_thyroxine
on_antithyroid_medication
sick
pregnant
thyroid_surgery
I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH_1
TSH_2
TSH_3
TSH_4
TSH_5
TSH_6
TSH_7
TSH_8
TSH_9
T3_measured
T3_1
T3_2
T3_3
T3_4
T3_5
T3_6
T3_7
T3_8
T3_9
TT4_measured
TT4_1
TT4_2
TT4_3
TT4_4
TT4_5
TT4_6
TT4_7
TT4_8
TT4_9
T4U_measured
T4U_1
T4U_2
T4U_3
T4U_4
T4U_5
T4U_6
T4U_7
T4U_8
T4U_9
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
FTI_5
FTI_6
FTI_7
FTI_8
FTI_9
TBG_measured
TBG
referral_source
Class
Test mode: 10-fold cross-validation

==== Classifier model (full training set) ====
J48 pruned tree
-----
```

## Cross Validation 15 folds

```
==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B10-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  78
              age_1
              age_2
              age_3
              age_4
              age_5
```

```

age_6
age_7
age_8
age_9
sex
on_thyroxine
query_on_thyroxine
on_antithyroid_medication
sick
pregnant
thyroid_surgery
I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH_1
TSH_2
TSH_3
TSH_4
TSH_5
TSH_6
TSH_7
TSH_8
TSH_9
T3_measured
T3_1
T3_2
T3_3
T3_4
T3_5
T3_6
T3_7
T3_8
T3_9
TT4_measured
TT4_1
TT4_2
TT4_3
TT4_4
TT4_5
TT4_6
TT4_7
TT4_8
TT4_9
T4U_measured
T4U_1
T4U_2
T4U_3
T4U_4
T4U_5
T4U_6
T4U_7
T4U_8
T4U_9
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
FTI_5
FTI_6
FTI_7
FTI_8
FTI_9
TBG_measured
TBG
referral_source
Class
Test mode: 15-fold cross-validation
==== Classifier model (full training set) ====
J48 pruned tree
-----
T3_1 = '(-inf-1.15]'
|   T3_measured = t
|   |   FTI_1 = '(-inf-78.5]'
|   |   |   sick = f
|   |   |   |   TSH_6 = '(-inf-1.75]': sick (13.92/2.59)
|   |   |   |   TSH_6 = '(1.75-inf)'
|   |   |   |   |   TSH_9 = '(-inf-6.15]'

```

```

|   |   |   |   |   age_4 = '(-inf-47.5]': sick (2.0)
|   |   |   |   age_4 = '(47.5-inf)'
|   |   |   |   age_7 = '(-inf-64.5]': negative (3.29/0.29)
|   |   |   |   age_7 = '(64.5-inf)': sick (5.71/2.29)
|   |   |   TSH_9 = '(6.15-inf)': negative (62.89/9.01)
|   |   |   sick = t: sick (3.29)
|   |   FTI_1 = '(78.5-inf)'
|   |   referral_source = SVHC: sick (9.0)
|   |   referral_source = other
|   |   |   TSH_9 = '(-inf-6.15]'
|   |   |   age_8 = '(-inf-70.5]': negative (23.95/2.0)
|   |   |   age_8 = '(70.5-inf)': sick (11.71/2.71)
|   |   |   TSH_9 = '(6.15-inf)': sick (6.46/0.46)
|   |   |   referral_source = SVI
|   |   |   FTI_9 = '(-inf-147.5]'
|   |   |   |   TT4_1 = '(-inf-72.5]'
|   |   |   |   age_2 = '(-inf-33.5]': negative (2.62)
|   |   |   |   age_2 = '(33.5-inf)': sick (24.87/4.24)
|   |   |   |   TT4_1 = '(72.5-inf)': sick (118.35)
|   |   |   FTI_9 = '(147.5-inf]'
|   |   |   |   TSH_1 = '(-inf-0.085]': negative (3.0)
|   |   |   |   TSH_1 = '(0.085-inf)'
|   |   |   |   TT4_8 = '(-inf-130.5]'
|   |   |   |   |   age_5 = '(-inf-54.5]': negative (3.34/1.17)
|   |   |   |   |   age_5 = '(54.5-inf)': sick (7.43/1.09)
|   |   |   |   |   TT4_8 = '(130.5-inf)': negative (6.17/1.17)
|   |   |   |   referral_source = STMW: sick (0.0)
|   |   |   |   referral_source = SVHD: sick (3.0)
|   T3_measured = f: negative (79.64/0.31)
T3_1 = '(1.15-inf)': negative (3381.36/24.69)

Number of Leaves : 21
Size of the tree : 38

Time taken to build model: 0.08 seconds

==== Stratified cross-validation ====
==== Summary ====

    Correctly Classified Instances      3689          97.7996 %
    Incorrectly Classified Instances     83           2.2004 %
    Kappa statistic                      0.8001
    Mean absolute error                  0.0305
    Root mean squared error              0.1376
    Relative absolute error              26.4388 %
    Root relative squared error         57.4062 %
    Total Number of Instances            3772

==== Detailed Accuracy By Class ====

          TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
          0.991     0.225     0.985     0.991     0.988     0.801     0.934     0.990     negative
          0.775     0.009     0.852     0.775     0.812     0.801     0.934     0.767     sick
Weighted Avg.     0.978     0.212     0.977     0.978     0.978     0.801     0.934     0.976

==== Confusion Matrix ====

      a     b  <-- classified as
3510   31 |     a = negative
      52  179 |     b = sick

```

## Percentage Split 50%

```

==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B10-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  78
             age_1
             age_2
             age_3
             age_4
             age_5
             age_6
             age_7
             age_8
             age_9
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication

```

```

sick
pregnant
thyroid_surgery
I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH_1
TSH_2
TSH_3
TSH_4
TSH_5
TSH_6
TSH_7
TSH_8
TSH_9
T3_measured
T3_1
T3_2
T3_3
T3_4
T3_5
T3_6
T3_7
T3_8
T3_9
TT4_measured
TT4_1
TT4_2
TT4_3
TT4_4
TT4_5
TT4_6
TT4_7
TT4_8
TT4_9
T4U_measured
T4U_1
T4U_2
T4U_3
T4U_4
T4U_5
T4U_6
T4U_7
T4U_8
T4U_9
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
FTI_5
FTI_6
FTI_7
FTI_8
FTI_9
TBG_measured
TBG
referral_source
Class
Test mode: split 50.0% train, remainder test

==== Classifier model (full training set) ====
J48 pruned tree
-----
T3_1 = '(-inf-1.15)'
| T3_measured = t
| | FTI_1 = '(-inf-78.5)'
| | | sick = f
| | | | TSH_6 = '(-inf-1.75)': sick (13.92/2.59)
| | | | TSH_6 = '(1.75-inf)'
| | | | | TSH_9 = '(-inf-6.15]'
| | | | | | age_4 = '(-inf-47.5)': sick (2.0)
| | | | | | age_4 = '(47.5-inf)'
| | | | | | | age_7 = '(-inf-64.5)': negative (3.29/0.29)
| | | | | | | age_7 = '(64.5-inf)': sick (5.71/2.29)
| | | | | | TSH_9 = '(6.15-inf)': negative (62.89/9.01)
| | | | sick = t: sick (3.29)
| | | FTI_1 = '(78.5-inf)'
| | | | referral_source = SVHC: sick (9.0)

```

```

|   |   |   referral_source = other
|   |   |   TSH_9 = '(-inf-6.15]'
|   |   |   |   age_8 = '(-inf-70.5]': negative (23.95/2.0)
|   |   |   |   age_8 = '(70.5-inf)': sick (11.71/2.71)
|   |   |   |   TSH_9 = '(6.15-inf)': sick (6.46/0.46)
|   |   |   referral_source = SVI
|   |   |   FTI_9 = '(-inf-147.5]'
|   |   |   |   TT4_1 = '(-inf-72.5]'
|   |   |   |   |   age_2 = '(-inf-33.5]': negative (2.62)
|   |   |   |   |   age_2 = '(33.5-inf)': sick (24.87/4.24)
|   |   |   |   |   TT4_1 = '(72.5-inf)': sick (118.35)
|   |   |   |   FTI_9 = '(147.5-inf)'
|   |   |   |   TSH_1 = '(-inf-0.085]': negative (3.0)
|   |   |   |   TSH_1 = '(0.085-inf)'
|   |   |   |   |   TT4_8 = '(-inf-130.5]'
|   |   |   |   |   |   age_5 = '(-inf-54.5]': negative (3.34/1.17)
|   |   |   |   |   |   age_5 = '(54.5-inf)': sick (7.43/1.09)
|   |   |   |   |   |   TT4_8 = '(130.5-inf)': negative (6.17/1.17)
|   |   |   |   |   referral_source = STMW: sick (0.0)
|   |   |   |   |   referral_source = SVHD: sick (3.0)
|   T3_measured = f: negative (79.64/0.31)
T3_1 = '(1.15-inf)': negative (3381.36/24.69)

Number of Leaves : 21
Size of the tree : 38

Time taken to build model: 0.09 seconds

==== Evaluation on test split ====

Time taken to test model on test split: 0 seconds

==== Summary ====

```

	1850	98.0912 %
Correctly Classified Instances	1850	98.0912 %
Incorrectly Classified Instances	36	1.9088 %
Kappa statistic	0.8202	
Mean absolute error	0.0282	
Root mean squared error	0.1289	
Relative absolute error	24.3965 %	
Root relative squared error	53.4221 %	
Total Number of Instances	1886	

```

==== Detailed Accuracy By Class ====

```

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
negative	0.996	0.248	0.984	0.996	0.990	0.825	0.937	0.991	negative
sick	0.752	0.004	0.926	0.752	0.830	0.825	0.937	0.832	sick
Weighted Avg.	0.981	0.233	0.980	0.981	0.980	0.825	0.937	0.982	

```

==== Confusion Matrix ====

```

a	b	--> classified as
1762	7	a = negative
29	88	b = sick

## Percentage Split 60%

```

==== Run information ====

```

Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2  
Relation: sick-weka.filters.unsupervised.attribute.Discretize-D-F-B10-M-1.0-Rfirst-last-precision6  
Instances: 3772  
Attributes: 78  
age\_1  
age\_2  
age\_3  
age\_4  
age\_5  
age\_6  
age\_7  
age\_8  
age\_9  
sex  
on\_thyroxine  
query\_on\_thyroxine  
on\_antithyroid\_medication  
sick  
pregnant  
thyroid\_surgery  
I131\_treatment  
query\_hypothyroid

```

query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH_1
TSH_2
TSH_3
TSH_4
TSH_5
TSH_6
TSH_7
TSH_8
TSH_9
T3_measured
T3_1
T3_2
T3_3
T3_4
T3_5
T3_6
T3_7
T3_8
T3_9
TT4_measured
TT4_1
TT4_2
TT4_3
TT4_4
TT4_5
TT4_6
TT4_7
TT4_8
TT4_9
T4U_measured
T4U_1
T4U_2
T4U_3
T4U_4
T4U_5
T4U_6
T4U_7
T4U_8
T4U_9
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
FTI_5
FTI_6
FTI_7
FTI_8
FTI_9
TBG_measured
TBG
referral_source
Class
Test mode: split 60.0% train, remainder test

==== Classifier model (full training set) ====

J48 pruned tree
-----
T3_1 = '(-inf-1.15)'
| T3_measured = t
| | FTI_1 = '(-inf-78.5)'
| | | sick = f
| | | | TSH_6 = '(-inf-1.75]': sick (13.92/2.59)
| | | | TSH_6 = '(1.75-inf]'
| | | | | TSH_9 = '(-inf-6.15]'
| | | | | | age_4 = '(-inf-47.5]': sick (2.0)
| | | | | | | age_4 = '(47.5-inf]'
| | | | | | | age_7 = '(-inf-64.5]': negative (3.29/0.29)
| | | | | | | age_7 = '(64.5-inf)': sick (5.71/2.29)
| | | | | | TSH_9 = '(6.15-inf)': negative (62.89/9.01)
| | | | sick = t: sick (3.29)
| | | | FTI_1 = '(78.5-inf)'
| | | | | referral_source = SVHC: sick (9.0)
| | | | | referral_source = other
| | | | | TSH_9 = '(-inf-6.15]'
| | | | | | age_8 = '(-inf-70.5]': negative (23.95/2.0)
| | | | | | age_8 = '(70.5-inf)': sick (11.71/2.71)
| | | | | TSH_9 = '(6.15-inf)': sick (6.46/0.46)

```

```

|   |   |   referral_source = SVI
|   |   |   |   FTI_9 = '(-inf-147.5]'
|   |   |   |   TT4_1 = '(-inf-72.5]'
|   |   |   |   |   age_2 = '(-inf-33.5]': negative (2.62)
|   |   |   |   |   age_2 = '(33.5-inf)': sick (24.87/4.24)
|   |   |   |   TT4_1 = '(72.5-inf)': sick (118.35)
|   |   |   |   FTI_9 = '(147.5-inf)'
|   |   |   |   TSH_1 = '(-inf-0.085]': negative (3.0)
|   |   |   |   TSH_1 = '(0.085-inf)'
|   |   |   |   |   TT4_8 = '(-inf-130.5]'
|   |   |   |   |   age_5 = '(-inf-54.5]': negative (3.34/1.17)
|   |   |   |   |   age_5 = '(54.5-inf)': sick (7.43/1.09)
|   |   |   |   |   TT4_8 = '(130.5-inf)': negative (6.17/1.17)
|   |   |   referral_source = STMW: sick (0.0)
|   |   |   referral_source = SVHD: sick (3.0)
|   T3_measured = f: negative (79.64/0.31)
T3_1 = '(1.15-inf)': negative (3381.36/24.69)

Number of Leaves : 21
Size of the tree : 38

Time taken to build model: 0.1 seconds
== Evaluation on test split ==

Time taken to test model on test split: 0 seconds
== Summary ==
Correctly Classified Instances      1478          97.9457 %
Incorrectly Classified Instances    31           2.0543 %
Kappa statistic                      0.7967
Mean absolute error                  0.029
Root mean squared error              0.1393
Relative absolute error              25.3118 %
Root relative squared error         59.14 %
Total Number of Instances           1509

== Detailed Accuracy By Class ==
            TP Rate  FP Rate  Precision  Recall   F-Measure  MCC     ROC Area  PRC Area  Class
               0.995   0.270    0.983    0.995    0.989    0.802    0.896    0.986    negative
               0.730   0.005    0.903    0.730    0.807    0.802    0.896    0.742    sick
Weighted Avg.       0.979   0.254    0.979    0.979    0.978    0.802    0.896    0.972

== Confusion Matrix ==
      a     b  <-- classified as
1413   7 |     a = negative
      24   65 |     b = sick

```

## Percentage Split 66%

```

== Run information ==
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B10-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  78
age_1
age_2
age_3
age_4
age_5
age_6
age_7
age_8
age_9
sex
on_thyroxine
query_on_thyroxine
on_antithyroid_medication
sick
pregnant
thyroid_surgery
t131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary

```

```

psych
TSH_measured
TSH_1
TSH_2
TSH_3
TSH_4
TSH_5
TSH_6
TSH_7
TSH_8
TSH_9
T3_measured
T3_1
T3_2
T3_3
T3_4
T3_5
T3_6
T3_7
T3_8
T3_9
TT4_measured
TT4_1
TT4_2
TT4_3
TT4_4
TT4_5
TT4_6
TT4_7
TT4_8
TT4_9
T4U_measured
T4U_1
T4U_2
T4U_3
T4U_4
T4U_5
T4U_6
T4U_7
T4U_8
T4U_9
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
FTI_5
FTI_6
FTI_7
FTI_8
FTI_9
TBG_measured
TBG
referral_source
Class
Test mode: split 66.0% train, remainder test

==== Classifier model (full training set) ====

J48 pruned tree
-----
T3_1 = '(-inf-1.15]'
| T3_measured = t
| | FTI_1 = '(-inf-78.5]'
| | | sick = f
| | | | TSH_6 = '(-inf-1.75]': sick (13.92/2.59)
| | | | TSH_6 = '(1.75-inf)'
| | | | TSH_9 = '(-inf-6.15]'
| | | | | age_4 = '(-inf-47.5]': sick (2.0)
| | | | | age_4 = '(47.5-inf)'
| | | | | | age_7 = '(-inf-64.5]': negative (3.29/0.29)
| | | | | | age_7 = '(64.5-inf)': sick (5.71/2.29)
| | | | | | TSH_9 = '(6.15-inf)': negative (62.89/9.01)
| | | | sick = t: sick (3.29)
| | FTI_1 = '(78.5-inf)'
| | | referral_source = SVHC: sick (9.0)
| | | referral_source = other
| | | | TSH_9 = '(-inf-6.15]'
| | | | | age_8 = '(-inf-70.5]': negative (23.95/2.0)
| | | | | age_8 = '(70.5-inf)': sick (11.71/2.71)
| | | | | TSH_9 = '(6.15-inf)': sick (6.46/0.46)
| | | | referral_source = SVI
| | | | | FTI_9 = '(-inf-147.5]'
| | | | | TT4_1 = '(-inf-72.5]'
| | | | | | age_2 = '(-inf-33.5]': negative (2.62)
| | | | | | age_2 = '(33.5-inf)': sick (24.87/4.24)

```

```

|   |   |   |   TT4_1 = '(72.5-inf)': sick (118.35)
|   |   |   |   FTI_9 = '(147.5-inf)'
|   |   |   |   TSH_1 = '(-inf-0.085)': negative (3.0)
|   |   |   |   TSH_1 = '(0.085-inf)'
|   |   |   |   |   TT4_8 = '(-inf-130.5]'
|   |   |   |   |   age_5 = '(-inf-54.5)': negative (3.34/1.17)
|   |   |   |   |   age_5 = '(54.5-inf)': sick (7.43/1.09)
|   |   |   |   |   TT4_8 = '(130.5-inf)': negative (6.17/1.17)
|   |   |   referral_source = STMW: sick (0.0)
|   |   |   referral_source = SVHD: sick (3.0)
|   T3_measured = f: negative (79.64/0.31)
T3_1 = '(1.15-inf)': negative (3381.36/24.69)

Number of Leaves : 21
Size of the tree : 38

Time taken to build model: 0.09 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

==== Summary ===

Correctly Classified Instances      1255           97.8939 %
Incorrectly Classified Instances    27            2.1061 %
Kappa statistic                   0.8026
Mean absolute error               0.0251
Root mean squared error          0.1348
Relative absolute error          22.0154 %
Root relative squared error     57.4181 %
Total Number of Instances        1282

==== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
0.991     0.213     0.987     0.991     0.989     0.803     0.894     0.984   negative
0.787     0.009     0.843     0.787     0.814     0.803     0.894     0.781   sick
Weighted Avg.  0.979     0.201     0.978     0.979     0.979     0.803     0.894     0.972

==== Confusion Matrix ===

      a      b  <-- classified as
1196   11 |     a = negative
      16   59 |     b = sick

```

## Leave one out fold

```

==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B10-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  78
              age_1
              age_2
              age_3
              age_4
              age_5
              age_6
              age_7
              age_8
              age_9
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH_1
              TSH_2
              TSH_3

```

```

TSH_4
TSH_5
TSH_6
TSH_7
TSH_8
TSH_9
T3_measured
T3_1
T3_2
T3_3
T3_4
T3_5
T3_6
T3_7
T3_8
T3_9
TT4_measured
TT4_1
TT4_2
TT4_3
TT4_4
TT4_5
TT4_6
TT4_7
TT4_8
TT4_9
T4U_measured
T4U_1
T4U_2
T4U_3
T4U_4
T4U_5
T4U_6
T4U_7
T4U_8
T4U_9
FTI_measured
FTI_1
FTI_2
FTI_3
FTI_4
FTI_5
FTI_6
FTI_7
FTI_8
FTI_9
TBG_measured
TBG
referral_source
Class
Test mode: 3772-fold cross-validation

==== Classifier model (full training set) ====
J48 pruned tree
-----
T3_1 = '(-inf-1.15)'
| T3_measured = t
| | FTI_1 = '(-inf-78.5)'
| | | sick = f
| | | | TSH_6 = '(-inf-1.75)': sick (13.92/2.59)
| | | | TSH_6 = '(1.75-inf)'
| | | | | TSH_9 = '(-inf-6.15)'
| | | | | age_4 = '(-inf-47.5)': sick (2.0)
| | | | | age_4 = '(47.5-inf)'
| | | | | | age_7 = '(-inf-64.5)': negative (3.29/0.29)
| | | | | | age_7 = '(64.5-inf)': sick (5.71/2.29)
| | | | | | TSH_9 = '(6.15-inf)': negative (62.89/9.01)
| | | | | sick = t: sick (3.29)
| | | | FTI_1 = '(78.5-inf)'
| | | | | referral_source = SVHC: sick (9.0)
| | | | | referral_source = other
| | | | | TSH_9 = '(-inf-6.15)'
| | | | | | age_8 = '(-inf-70.5)': negative (23.95/2.0)
| | | | | | age_8 = '(70.5-inf)': sick (11.71/2.71)
| | | | | | TSH_9 = '(6.15-inf)': sick (6.46/0.46)
| | | | | | referral_source = SVI
| | | | | | FTI_9 = '(-inf-147.5)'
| | | | | | | TT4_1 = '(-inf-72.5]'
| | | | | | | | age_2 = '(-inf-33.5)': negative (2.62)
| | | | | | | | age_2 = '(33.5-inf)': sick (24.87/4.24)
| | | | | | | | TT4_1 = '(72.5-inf)': sick (118.35)
| | | | | | | | FTI_9 = '(147.5-inf)'
| | | | | | | | TSH_1 = '(-inf-0.085)': negative (3.0)
| | | | | | | | TSH_1 = '(0.085-inf)'
| | | | | | | | TT4_8 = '(-inf-130.5)'

```

```

|   |   |   |   |   |   age_5 = '(-inf-54.5]': negative (3.34/1.17)
|   |   |   |   |   age_5 = '(54.5-inf)': sick (7.43/1.09)
|   |   |   |   TT4_8 = '(130.5-inf)': negative (6.17/1.17)
|   |   |   referral_source = STMW: sick (0.0)
|   |   |   referral_source = SVHD: sick (3.0)
|   T3_measured = f: negative (79.64/0.31)
T3_1 = '(1.15-inf)': negative (3381.36/24.69)

Number of Leaves : 21

Size of the tree : 38

Time taken to build model: 0.1 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3692          97.8791 %
Incorrectly Classified Instances    80           2.1209 %
Kappa statistic                   0.807
Mean absolute error                0.0318
Root mean squared error            0.1384
Relative absolute error             27.5845 %
Root relative squared error       57.7141 %
Total Number of Instances          3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC   ROC Area   PRC Area   Class
0.992     0.221     0.986     0.992     0.989     0.808     0.905     0.988   negative
0.779     0.008     0.861     0.779     0.818     0.808     0.905     0.833   sick
Weighted Avg.   0.979     0.208     0.978     0.979     0.978     0.808     0.905     0.979

==== Confusion Matrix ====

      a     b  <-- classified as
3512   29 |   a = negative
      51  180 |   b = sick

```

#### 4.2.4 40 bins

#### Cross Validation 5 folds

```

==== Run information ====

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B40-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  258
              [list of attributes omitted]
Test mode:   5-fold cross-validation

==== Classifier model (full training set) ====

J48 pruned tree
-----

T3_4 = '(-inf-1.15]'
|   T3_measured = t
|   |   TT4_1 = '(-inf-46.5]'
|   |   |   TSH_38 = '(-inf-12.55]'
|   |   |   T4U_1 = '(-inf-0.685)': negative (7.64/1.0)
|   |   |   T4U_1 = '(0.685-inf)': sick (4.36/0.36)
|   |   |   TSH_38 = '(12.55-inf)': negative (47.0/1.0)
|   |   TT4_1 = '(46.5-inf)'
|   |   |   referral_source = SVHC: sick (9.0)
|   |   |   referral_source = other
|   |   |   sick = f
|   |   |   |   age_33 = '(-inf-70.5]'
|   |   |   |   TSH_34 = '(-inf-4.05)': negative (26.4/1.0)
|   |   |   |   TSH_34 = '(4.05-inf)': sick (6.6/1.6)
|   |   |   |   age_33 = '(70.5-inf)'
|   |   |   |   FTI_2 = '(-inf-68.5)': negative (2.12)
|   |   |   |   FTI_2 = '(68.5-inf)': sick (15.88/2.88)
|   |   |   |   sick = t: sick (3.0)
|   |   |   referral_source = SVI
|   |   |   FTI_37 = '(-inf-153.5]'
|   |   |   |   TSH_38 = '(-inf-12.55)': sick (163.32/3.0)
|   |   |   |   TSH_38 = '(12.55-inf)': negative (8.05/2.05)
|   |   |   FTI_37 = '(153.5-inf)'
|   |   |   |   TT4_20 = '(-inf-104.5)': sick (3.31/1.0)
|   |   |   |   TT4_20 = '(104.5-inf)': negative (11.31/1.31)

```

```

|   |   |   referral_source = STMW: sick (0.0)
|   |   |   referral_source = SVHD: sick (3.0)
|   T3_measured = f: negative (79.64/0.31)
T3_4 = '(1.15-inf)'
|   FTI_38 = '(-inf-165.5)': negative (3203.91/8.69)
|   FTI_38 = '(165.5-inf)'
|   |   T3_19 = '(-inf-2.65)'
|   |   |   on_thyroxine = f
|   |   |   |   TT4_36 = '(-inf-150.5)': negative (19.73)
|   |   |   |   TT4_36 = '(150.5-inf)'
|   |   |   |   T3_measured = t
|   |   |   |   |   TSH_8 = '(-inf-0.285)': sick (17.55/1.55)
|   |   |   |   |   TSH_8 = '(0.285-inf)': negative (7.37)
|   |   |   |   T3_measured = f: negative (8.36)
|   |   |   |   on_thyroxine = t: negative (30.19)
|   |   T3_19 = '(2.65-inf)': negative (94.24)

Number of Leaves : 23
Size of the tree : 42

Time taken to build model: 0.35 seconds

==== Stratified cross-validation ====
==== Summary ====

    Correctly Classified Instances      3727           98.807 %
    Incorrectly Classified Instances     45            1.193 %
    Kappa statistic                   0.8939
    Mean absolute error               0.0166
    Root mean squared error          0.1062
    Relative absolute error          14.4259 %
    Root relative squared error      44.3014 %
    Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC     ROC Area   PRC Area   Class
      0.995     0.121     0.992     0.995     0.994     0.894     0.943     0.989   negative
      0.879     0.005     0.923     0.879     0.900     0.894     0.943     0.842   sick
Weighted Avg.       0.988     0.114     0.988     0.988     0.988     0.894     0.943     0.980

==== Confusion Matrix ====

      a     b  <-- classified as
3524   17 |     a = negative
      28  203 |     b = sick

```

## Cross Validation 10 folds

```

==== Run information ====

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B40-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  258
              [list of attributes omitted]
Test mode:   10-fold cross-validation

==== Classifier model (full training set) ====

J48 pruned tree
-----

T3_4 = '(-inf-1.15'
|   T3_measured = t
|   |   TT4_1 = '(-inf-46.5'
|   |   |   TSH_38 = '(-inf-12.55'
|   |   |   |   T4U_1 = '(-inf-0.685)': negative (7.64/1.0)
|   |   |   |   T4U_1 = '(0.685-inf)': sick (4.36/0.36)
|   |   |   TSH_38 = '(12.55-inf)': negative (47.0/1.0)
|   |   TT4_1 = '(46.5-inf'
|   |   |   referral_source = SVHC: sick (9.0)
|   |   |   referral_source = other
|   |   |   |   sick = f
|   |   |   |   |   age_33 = '(-inf-70.5'
|   |   |   |   |   |   TSH_34 = '(-inf-4.05)': negative (26.4/1.0)
|   |   |   |   |   |   TSH_34 = '(4.05-inf)': sick (6.6/1.6)
|   |   |   |   |   age_33 = '(70.5-inf'
|   |   |   |   |   |   FTI_2 = '(-inf-68.5)': negative (2.12)
|   |   |   |   |   |   FTI_2 = '(68.5-inf)': sick (15.88/2.88)
|   |   |   |   |   sick = t: sick (3.0)
|   |   |   referral_source = SVI

```

```

|   |   |   |   FTI_37 = '(-inf-153.5)'
|   |   |   |   TSH_38 = '(-inf-12.55)': sick (163.32/3.0)
|   |   |   |   TSH_38 = '(12.55-inf)': negative (8.05/2.05)
|   |   |   FTI_37 = '(153.5-inf)'
|   |   |   |   TT4_20 = '(-inf-104.5)': sick (3.31/1.0)
|   |   |   |   TT4_20 = '(104.5-inf)': negative (11.31/1.31)
|   |   |   referral_source = STMW: sick (0.0)
|   |   |   referral_source = SVHD: sick (3.0)
|   T3_measured = f: negative (79.64/0.31)
T3_4 = '(1.15-inf'
|   FTI_38 = '(-inf-165.5)': negative (3203.91/8.69)
|   FTI_38 = '(165.5-inf'
|   |   T3_19 = '(-inf-2.65'
|   |   |   on_thyroxine = f
|   |   |   TT4_36 = '(-inf-150.5)': negative (19.73)
|   |   |   TT4_36 = '(150.5-inf'
|   |   |   |   T3_measured = t
|   |   |   |   |   TSH_8 = '(-inf-0.285)': sick (17.55/1.55)
|   |   |   |   |   TSH_8 = '(0.285-inf)': negative (7.37)
|   |   |   |   T3_measured = f: negative (8.36)
|   |   |   |   on_thyroxine = t: negative (30.19)
|   |   |   T3_19 = '(2.65-inf)': negative (94.24)

Number of Leaves : 23
Size of the tree : 42

Time taken to build model: 0.3 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3727          98.807 %
Incorrectly Classified Instances    45           1.193 %
Kappa statistic                   0.8934
Mean absolute error               0.0168
Root mean squared error          0.1061
Relative absolute error          14.5952 %
Root relative squared error     44.2315 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
0.995    0.126    0.992    0.995    0.994    0.894    0.944    0.989    negative
0.874    0.005    0.927    0.874    0.900    0.894    0.944    0.849    sick
Weighted Avg.  0.988    0.118    0.988    0.988    0.988    0.894    0.944    0.981

==== Confusion Matrix ====

      a      b  <-- classified as
3525   16 |      a = negative
      29  202 |      b = sick

```

## Cross Validation 15 folds

```

==== Run information ====

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B40-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  258
            [list of attributes omitted]
Test mode:   15-fold cross-validation

==== Classifier model (full training set) ====

J48 pruned tree
-----
T3_4 = '(-inf-1.15'
|   T3_measured = t
|   |   TT4_1 = '(-inf-46.5'
|   |   |   TSH_38 = '(-inf-12.55'
|   |   |   T4U_1 = '(-inf-0.685)': negative (7.64/1.0)
|   |   |   T4U_1 = '(0.685-inf)': sick (4.36/0.36)
|   |   |   TSH_38 = '(12.55-inf)': negative (47.0/1.0)
|   |   TT4_1 = '(46.5-inf'
|   |   |   referral_source = SVHC: sick (9.0)
|   |   |   referral_source = other
|   |   |   sick = f
|   |   |   |   age_33 = '(-inf-70.5'
|   |   |   |   TSH_34 = '(-inf-4.05)': negative (26.4/1.0)

```

```

|   |   |   |   |   TSH_34 = '(4.05-inf)': sick (6.6/1.6)
|   |   |   |   age_33 = '(70.5-inf)'
|   |   |   |   |   FTI_2 = '(-inf-68.5)': negative (2.12)
|   |   |   |   |   FTI_2 = '(68.5-inf)': sick (15.88/2.88)
|   |   |   |   sick = t: sick (3.0)
|   |   |   referral_source = SVI
|   |   |   |   FTI_37 = '(-inf-153.5]'
|   |   |   |   |   TSH_38 = '(-inf-12.55)': sick (163.32/3.0)
|   |   |   |   |   TSH_38 = '(12.55-inf)': negative (8.05/2.05)
|   |   |   |   FTI_37 = '(153.5-inf)'
|   |   |   |   |   TT4_20 = '(-inf-104.5)': sick (3.31/1.0)
|   |   |   |   |   TT4_20 = '(104.5-inf)': negative (11.31/1.31)
|   |   |   referral_source = STMW: sick (0.0)
|   |   |   referral_source = SVHD: sick (3.0)
|   T3_measured = f: negative (79.64/0.31)
T3_4 = '(1.15-inf)'
|   FTI_38 = '(-inf-165.5)': negative (3203.91/8.69)
|   FTI_38 = '(165.5-inf)'
|   |   T3_19 = '(-inf-2.65)'
|   |   |   on_thyroxine = f
|   |   |   |   TT4_36 = '(-inf-150.5)': negative (19.73)
|   |   |   |   TT4_36 = '(150.5-inf)'
|   |   |   |   |   T3_measured = t
|   |   |   |   |   |   TSH_8 = '(-inf-0.285)': sick (17.55/1.55)
|   |   |   |   |   |   TSH_8 = '(0.285-inf)': negative (7.37)
|   |   |   |   |   T3_measured = f: negative (8.36)
|   |   |   |   on_thyroxine = t: negative (30.19)
|   |   T3_19 = '(2.65-inf)': negative (94.24)

```

Number of Leaves : 23

Size of the tree : 42

Time taken to build model: 0.32 seconds

==== Stratified cross-validation ====  
==== Summary ===

Correctly Classified Instances	3725	98.754 %
Incorrectly Classified Instances	47	1.246 %
Kappa statistic	0.891	
Mean absolute error	0.0173	
Root mean squared error	0.1074	
Relative absolute error	15.0167 %	
Root relative squared error	44.8042 %	
Total Number of Instances	3772	

==== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
0.994	0.108	0.993	0.994	0.993	0.891	0.932	0.986		negative
0.892	0.006	0.904	0.892	0.898	0.891	0.932	0.810		sick
Weighted Avg.	0.988	0.102	0.987	0.988	0.988	0.891	0.932	0.975	

==== Confusion Matrix ===

a	b	<-- classified as
3519	22	a = negative
25	206	b = sick

## Percentage Split 50%

==== Run information ===

Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2  
 Relation: sick-weka.filters.unsupervised.attribute.Discretize-D-F-B40-M-1.0-Rfirst-last-precision6  
 Instances: 3772  
 Attributes: 258  
 [list of attributes omitted]  
 Test mode: split 50.0% train, remainder test

==== Classifier model (full training set) ===

J48 pruned tree

```

T3_4 = '(-inf-1.15]'
|   T3_measured = t
|   |   TT4_1 = '(-inf-46.5]'
|   |   |   TSH_38 = '(-inf-12.55]'
|   |   |   |   T4U_1 = '(-inf-0.685)': negative (7.64/1.0)
|   |   |   |   T4U_1 = '(0.685-inf)': sick (4.36/0.36)
|   |   |   TSH_38 = '(12.55-inf)': negative (47.0/1.0)

```

```

|   |   TT4_1 = '(46.5-inf)'
|   |   | referral_source = SVHC: sick (9.0)
|   |   | referral_source = other
|   |   |   | sick = f
|   |   |   |   | age_33 = '(-inf-70.5]'
|   |   |   |   |   | TSH_34 = '(-inf-4.05]': negative (26.4/1.0)
|   |   |   |   |   | TSH_34 = '(4.05-inf)': sick (6.6/1.6)
|   |   |   |   |   | age_33 = '(70.5-inf)'
|   |   |   |   |   | FTI_2 = '(-inf-68.5]': negative (2.12)
|   |   |   |   |   | FTI_2 = '(68.5-inf)': sick (15.88/2.88)
|   |   |   |   |   | sick = t: sick (3.0)
|   |   |   |   |   | referral_source = SVI
|   |   |   |   |   | FTI_37 = '(-inf-153.5]'
|   |   |   |   |   |   | TSH_38 = '(-inf-12.55]': sick (163.32/3.0)
|   |   |   |   |   |   | TSH_38 = '(12.55-inf)': negative (8.05/2.05)
|   |   |   |   |   |   | FTI_37 = '(153.5-inf)'
|   |   |   |   |   |   | TT4_20 = '(-inf-104.5]': sick (3.31/1.0)
|   |   |   |   |   |   | TT4_20 = '(104.5-inf)': negative (11.31/1.31)
|   |   |   |   |   |   | referral_source = STMW: sick (0.0)
|   |   |   |   |   |   | referral_source = SVHD: sick (3.0)
|   |   |   |   |   |   | T3_measured = f: negative (79.64/0.31)
T3_4 = '(1.15-inf'
|   |   |   |   |   |   | FTI_38 = '(-inf-165.5]': negative (3203.91/8.69)
|   |   |   |   |   |   | FTI_38 = '(165.5-inf)'
|   |   |   |   |   |   | T3_19 = '(-inf-2.65]'
|   |   |   |   |   |   |   | on_thyroxine = f
|   |   |   |   |   |   |   | TT4_36 = '(-inf-150.5]': negative (19.73)
|   |   |   |   |   |   |   | TT4_36 = '(150.5-inf)'
|   |   |   |   |   |   |   |   | T3_measured = t
|   |   |   |   |   |   |   |   |   | TSH_8 = '(-inf-0.285]': sick (17.55/1.55)
|   |   |   |   |   |   |   |   |   | TSH_8 = '(0.285-inf)': negative (7.37)
|   |   |   |   |   |   |   |   |   | T3_measured = f: negative (8.36)
|   |   |   |   |   |   |   |   |   | on_thyroxine = t: negative (30.19)
|   |   |   |   |   |   |   |   |   | T3_19 = '(2.65-inf)': negative (94.24)

Number of Leaves : 23
Size of the tree : 42

Time taken to build model: 0.31 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0.02 seconds

==== Summary ===

    Correctly Classified Instances      1863           98.7805 %
    Incorrectly Classified Instances     23            1.2195 %
    Kappa statistic                      0.8846
    Mean absolute error                  0.0201
    Root mean squared error              0.1079
    Relative absolute error              17.4049 %
    Root relative squared error          44.7451 %
    Total Number of Instances           1886

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      1.000     0.197     0.987     1.000     0.994     0.891     0.924     0.989   negative
      0.803     0.000     1.000     0.803     0.891     0.891     0.924     0.868   sick
Weighted Avg.     0.988     0.184     0.988     0.988     0.987     0.891     0.924     0.981

==== Confusion Matrix ===

      a     b  <-- classified as
1769   0 |   a = negative
      23   94 |   b = sick

```

## Percentage Split 60%

```

==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B40-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  258
              [list of attributes omitted]
Test mode:   split 60.0% train, remainder test

==== Classifier model (full training set) ===

J48 pruned tree

```

```

-----
T3_4 = '(-inf-1.15)'
| T3_measured = t
| | TT4_1 = '(-inf-46.5)'
| | | TSH_38 = '(-inf-12.55)'
| | | | T4U_1 = '(-inf-0.685)': negative (7.64/1.0)
| | | | T4U_1 = '(0.685-inf)': sick (4.36/0.36)
| | | | TSH_38 = '(12.55-inf)': negative (47.0/1.0)
| | | | TT4_1 = '(46.5-inf)'
| | | | referral_source = SVHC: sick (9.0)
| | | | referral_source = other
| | | | | sick = f
| | | | | age_33 = '(-inf-70.5)'
| | | | | TSH_34 = '(-inf-4.05)': negative (26.4/1.0)
| | | | | TSH_34 = '(4.05-inf)': sick (6.6/1.6)
| | | | | age_33 = '(70.5-inf)'
| | | | | | FTI_2 = '(-inf-68.5)': negative (2.12)
| | | | | | FTI_2 = '(68.5-inf)': sick (15.88/2.88)
| | | | | sick = t: sick (3.0)
| | | | | referral_source = SVI
| | | | | FTI_37 = '(-inf-153.5)'
| | | | | TSH_38 = '(-inf-12.55)': sick (163.32/3.0)
| | | | | TSH_38 = '(12.55-inf)': negative (8.05/2.05)
| | | | | FTI_37 = '(153.5-inf)'
| | | | | | TT4_20 = '(-inf-104.5)': sick (3.31/1.0)
| | | | | | TT4_20 = '(104.5-inf)': negative (11.31/1.31)
| | | | | referral_source = STMW: sick (0.0)
| | | | | referral_source = SVHD: sick (3.0)
| | T3_measured = f: negative (79.64/0.31)
T3_4 = '(1.15-inf)'
| FTI_38 = '(-inf-165.5)': negative (3203.91/8.69)
| FTI_38 = '(165.5-inf)'
| | T3_19 = '(-inf-2.65)'
| | | on_thyroxine = f
| | | | TT4_36 = '(-inf-150.5)': negative (19.73)
| | | | TT4_36 = '(150.5-inf)'
| | | | | T3_measured = t
| | | | | | TSH_8 = '(-inf-0.285)': sick (17.55/1.55)
| | | | | | TSH_8 = '(0.285-inf)': negative (7.37)
| | | | | T3_measured = f: negative (8.36)
| | | | | on_thyroxine = t: negative (30.19)
| | | | | T3_19 = '(2.65-inf)': negative (94.24)

Number of Leaves : 23
Size of the tree : 42

Time taken to build model: 0.32 seconds
==== Evaluation on test split ====
Time taken to test model on test split: 0.03 seconds
==== Summary ====
Correctly Classified Instances      1488      98.6083 %
Incorrectly Classified Instances     21       1.3917 %
Kappa statistic                      0.8607
Mean absolute error                  0.0225
Root mean squared error              0.1153
Relative absolute error              19.5875 %
Root relative squared error         48.9526 %
Total Number of Instances           1509

==== Detailed Accuracy By Class ====


|               | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC   | ROC Area | PRC Area | Class    |
|---------------|---------|---------|-----------|--------|-----------|-------|----------|----------|----------|
| 0.999         | 0.225   | 0.986   | 0.999     | 0.993  | 0.868     | 0.910 | 0.988    |          | negative |
| 0.775         | 0.001   | 0.986   | 0.775     | 0.868  | 0.868     | 0.910 | 0.852    |          | sick     |
| Weighted Avg. | 0.986   | 0.212   | 0.986     | 0.986  | 0.985     | 0.868 | 0.910    | 0.980    |          |


==== Confusion Matrix ====


| a    | b  | <-- classified as |
|------|----|-------------------|
| 1419 | 1  | a = negative      |
| 20   | 69 | b = sick          |


```

## Percentage Split 66%

```

==== Run information ====
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2

```

```

Relation: sick-weka.filters.unsupervised.attribute.Discretize-D-F-B40-M-1.0-Rfirst-last-precision6
Instances: 3772
Attributes: 258
[list of attributes omitted]
Test mode: split 66.0% train, remainder test

==== Classifier model (full training set) ===

J48 pruned tree
-----

T3_4 = '(-inf-1.15)'
| T3_measured = t
| | TT4_1 = '(-inf-46.5)'
| | | TSH_38 = '(-inf-12.55)'
| | | | T4U_1 = '(-inf-0.685)': negative (7.64/1.0)
| | | | T4U_1 = '(0.685-inf)': sick (4.36/0.36)
| | | | TSH_38 = '(12.55-inf)': negative (47.0/1.0)
| | | | TT4_1 = '(46.5-inf)'
| | | | referral_source = SVHC: sick (9.0)
| | | | referral_source = other
| | | | sick = f
| | | | | age_33 = '(-inf-70.5)'
| | | | | TSH_34 = '(-inf-4.05)': negative (26.4/1.0)
| | | | | TSH_34 = '(4.05-inf)': sick (6.6/1.6)
| | | | | age_33 = '(70.5-inf)'
| | | | | FTI_2 = '(-inf-68.5)': negative (2.12)
| | | | | FTI_2 = '(68.5-inf)': sick (15.88/2.88)
| | | | | sick = t: sick (3.0)
| | | | | referral_source = SVI
| | | | | FTI_37 = '(-inf-153.5)'
| | | | | TSH_38 = '(-inf-12.55)': sick (163.32/3.0)
| | | | | TSH_38 = '(12.55-inf)': negative (8.05/2.05)
| | | | | FTI_37 = '(153.5-inf)'
| | | | | TT4_20 = '(-inf-104.5)': sick (3.31/1.0)
| | | | | TT4_20 = '(104.5-inf)': negative (11.31/1.31)
| | | | | referral_source = STMW: sick (0.0)
| | | | | referral_source = SVHD: sick (3.0)
| | | | | T3_measured = f: negative (79.64/0.31)
T3_4 = '(1.15-inf)'
| FTI_38 = '(-inf-165.5)': negative (3203.91/8.69)
| FTI_38 = '(165.5-inf)'
| | T3_19 = '(-inf-2.65)'
| | | on_thyroxine = f
| | | | TT4_36 = '(-inf-150.5)': negative (19.73)
| | | | TT4_36 = '(150.5-inf)'
| | | | | T3_measured = t
| | | | | TSH_8 = '(-inf-0.285)': sick (17.55/1.55)
| | | | | TSH_8 = '(0.285-inf)': negative (7.37)
| | | | | T3_measured = f: negative (8.36)
| | | | | on_thyroxine = t: negative (30.19)
| | | | | T3_19 = '(2.65-inf)': negative (94.24)

Number of Leaves : 23
Size of the tree : 42

Time taken to build model: 0.35 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0.03 seconds

==== Summary ===

    Correctly Classified Instances      1269          98.986 %
    Incorrectly Classified Instances     13           1.014 %
    Kappa statistic                      0.905
    Mean absolute error                  0.0156
    Root mean squared error              0.0997
    Relative absolute error              13.6583 %
    Root relative squared error         42.4824 %
    Total Number of Instances            1282

==== Detailed Accuracy By Class ===

          TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
          0.997     0.120     0.993     0.997     0.995     0.906     0.934     0.990     negative
          0.880     0.003     0.943     0.880     0.910     0.906     0.934     0.873     sick
Weighted Avg.   0.990     0.113     0.990     0.990     0.990     0.906     0.934     0.983

==== Confusion Matrix ===

    a     b    -- classified as
1203   4 |    a = negative
      9   66 |    b = sick

```

## Leave one out fold

```

==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Discretize-D-F-B40-M-1.0-Rfirst-last-precision6
Instances:   3772
Attributes:  258
              [list of attributes omitted]
Test mode:   3772-fold cross-validation

==== Classifier model (full training set) ===

J48 pruned tree
-----

T3_4 = '(-inf-1.15)'
| T3_measured = t
| | TT4_1 = '(-inf-46.5)'
| | | TSH_38 = '(-inf-12.55)'
| | | | T4U_1 = '(-inf-0.685)': negative (7.64/1.0)
| | | | T4U_1 = '(0.685-inf)': sick (4.36/0.36)
| | | | TSH_38 = '(12.55-inf)': negative (47.0/1.0)
| | | | TT4_1 = '(46.5-inf)'
| | | | referral_source = SVHC: sick (9.0)
| | | | referral_source = other
| | | | | sick = f
| | | | | age_33 = '(-inf-70.5)'
| | | | | TSH_34 = '(-inf-4.05)': negative (26.4/1.0)
| | | | | TSH_34 = '(4.05-inf)': sick (6.6/1.6)
| | | | | age_33 = '(70.5-inf)'
| | | | | FTI_2 = '(-inf-68.5)': negative (2.12)
| | | | | FTI_2 = '(68.5-inf)': sick (15.88/2.88)
| | | | | sick = t: sick (3.0)
| | | | | referral_source = SVI
| | | | | FTI_37 = '(-inf-153.5)'
| | | | | TSH_38 = '(-inf-12.55)': sick (163.32/3.0)
| | | | | TSH_38 = '(12.55-inf)': negative (8.05/2.05)
| | | | | FTI_37 = '(153.5-inf)'
| | | | | TT4_20 = '(-inf-104.5)': sick (3.31/1.0)
| | | | | TT4_20 = '(104.5-inf)': negative (11.31/1.31)
| | | | | referral_source = STMW: sick (0.0)
| | | | | referral_source = SVHD: sick (3.0)
| T3_measured = f: negative (79.64/0.31)
T3_4 = '(1.15-inf)'
| FTI_38 = '(-inf-165.5)': negative (3203.91/8.69)
| FTI_38 = '(165.5-inf)'
| | T3_19 = '(-inf-2.65)'
| | | on_thyroxine = f
| | | | TT4_36 = '(-inf-150.5)': negative (19.73)
| | | | TT4_36 = '(150.5-inf)'
| | | | | T3_measured = t
| | | | | TSH_8 = '(-inf-0.285)': sick (17.55/1.55)
| | | | | TSH_8 = '(0.285-inf)': negative (7.37)
| | | | | T3_measured = f: negative (8.36)
| | | | | on_thyroxine = t: negative (30.19)
| | | | | T3_19 = '(2.65-inf)': negative (94.24)

Number of Leaves : 23
Size of the tree : 42

Time taken to build model: 0.31 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3732          98.9396 %
Incorrectly Classified Instances     40           1.0604 %
Kappa statistic                      0.9059
Mean absolute error                  0.0163
Root mean squared error                0.1005
Relative absolute error                 14.1348 %
Root relative squared error            41.884  %
Total Number of Instances             3772

==== Detailed Accuracy By Class ====

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
0.996     0.108     0.993     0.996     0.994     0.906     0.939     0.988  negative
0.892     0.004     0.932     0.892     0.912     0.906     0.939     0.880      sick
Weighted Avg.    0.989     0.102     0.989     0.989     0.989     0.906     0.939     0.981

```

```
==== Confusion Matrix ====
      a     b  <-- classified as
3526   15 |     a = negative
      25  206 |     b = sick
```

#### 4.2.5 Conclusion

C4.5							
Preprocessing method	Cross Validation			Percentage Split			Leave One Out Fold
	5 folds	10 folds	15 folds	50.00%	60.00%	66.00%	
With Original Dataset	98.4358%	98.8070%	98.6479%	98.3563%	98.6083%	98.0499%	98.5154%
Equal frequency binning	40 bins	98.8070%	98.8070%	98.7540%	98.7805%	98.6083%	98.9396%
	10 bins	98.1442%	98.0382%	97.7996%	98.0912%	97.9457%	97.8939%
	5 bins	94.6978%	94.6713%	94.6978%	94.5387%	95.0298%	94.5398%
	2 bins	93.8759%	93.8759%	93.8759%	93.7964%	94.1021%	94.1498%
							93.8759%

Discretizing with equal frequency binning and 40 bins shows impressive results winning in 7 out 7 tests which shows that C4.5 can benefit from this preprocess step. other bins didn't show any improvements in performance.

### 4.3 Supervised discretization

#### Cross Validation 5 folds

```
==== Run information ====
Scheme:      weka.classifiers.meta.FilteredClassifier -F "weka.filters.supervised.attribute.Discretize -D -R first-last -precision 6" -S 1 -W
Relation:    sick
Instances:   3772
Attributes:  30
             age
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             i131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary
             psych
             TSH_measured
             TSH
             T3_measured
             T3
             TT4_measured
             TT4
             T4U_measured
             T4U
             FTI_measured
             FTI
             TBG_measured
             TBG
             referral_source
             Class
Test mode:   5-fold cross-validation

==== Classifier model (full training set) ====
FilteredClassifier using weka.classifiers.trees.J48 -C 0.25 -M 2 on data filtered through weka.filters.supervised.attribute.Discretize -D -R f
First
@relation sick-weka.filters.supervised.attribute.Discretize-D-Rfirst-last-precision6

@attribute age_1 (''(-inf-43.5)'',''(43.5-inf)'')
@attribute age_2 (''(-inf-69.5)'',''(69.5-inf)'')
@attribute sex {F,M}
```

```

@attribute on_thyroxine {f,t}
@attribute query_on_thyroxine {f,t}
@attribute on_antithyroid_medication {f,t}
@attribute sick {f,t}
@attribute pregnant {f,t}
@attribute thyroid_surgery {f,t}
@attribute I131_treatment {f,t}
@attribute query_hypothyroid {f,t}
@attribute query_hyperthyroid {f,t}
@attribute lithium {f,t}
@attribute goitre {f,t}
@attribute tumor {f,t}
@attribute hypopituitary {f,t}
@attribute psych {f,t}
@attribute TSH_measured {t,f}
@attribute TSH {'\All\'}
@attribute T3_measured {t,f}
@attribute T3_1 {'\(-inf-1.15)\', \'(1.15-inf)\'}
@attribute TT4_measured {t,f}
@attribute TT4_1 {'\(-inf-87.5)\', \'(87.5-inf)\'}
@attribute T4U_measured {t,f}
@attribute T4U_1 {'\(-inf-0.775)\', \'(0.775-inf)\'}
@attribute T4U_2 {'\(-inf-0.885)\', \'(0.885-inf)\'}
@attribute T4U_3 {'\(-inf-1.145)\', \'(1.145-inf)\'}
@attribute FTI_measured {t,f}
@attribute FTI {'\All\'}
@attribute TBG_measured {f}
@attribute TBG {'\All\'}
@attribute referral_source {SVHC,other,SVI,STMW,SVHD}
@attribute Class {negative,sick}

@data

Classifier Model
J48 pruned tree
-----
T3_1 = '(-inf-1.15)'
| T3_measured = t
| | referral_source = SVHC: sick (9.0)
| | referral_source = other
| | | sick = f
| | | | age_2 = '(-inf-69.5)'
| | | | | on_thyroxine = f: negative (50.0/3.0)
| | | | | on_thyroxine = t
| | | | | | age_1 = '(-inf-43.5)': sick (2.0)
| | | | | | age_1 = '(43.5-inf)': negative (5.0/1.0)
| | | | | age_2 = '(69.5-inf)'
| | | | | | T4U_measured = t: sick (19.0/6.0)
| | | | | | T4U_measured = f: negative (2.0)
| | | | | sick = t: sick (3.0)
| | | | | referral_source = SVI
| | | | | sick = f
| | | | | | TT4_1 = '(-inf-87.5)'
| | | | | | T4U_measured = t
| | | | | | | T4U_2 = '(-inf-0.885)'
| | | | | | | | query_hypothyroid = f: sick (78.0/16.0)
| | | | | | | | query_hypothyroid = t
| | | | | | | | | age_2 = '(-inf-69.5)': sick (3.0)
| | | | | | | | | age_2 = '(69.5-inf)': negative (5.0/1.0)
| | | | | | | | | T4U_2 = '(0.885-inf)'
| | | | | | | | | age_1 = '(-inf-43.5)': sick (2.0)
| | | | | | | | | age_1 = '(43.5-inf)'
| | | | | | | | | | query_hypothyroid = f: negative (14.0/2.0)
| | | | | | | | | | query_hypothyroid = t: sick (2.0)
| | | | | | | | | T4U_measured = f: negative (4.0)
| | | | | | | | | TT4_1 = '(87.5-inf)': sick (90.0/10.0)
| | | | | | | | | sick = t: sick (20.0)
| | | | | | | | | referral_source = STMW: sick (0.0)
| | | | | | | | | referral_source = SVHD: sick (3.0)
| | | | | | T3_measured = f: negative (79.64/0.31)
T3_1 = '(1.15-inf)': negative (3381.36/24.69)

Number of Leaves : 20
Size of the tree : 36

Time taken to build model: 0.46 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3695          97.9586 %
Incorrectly Classified Instances     77           2.0414 %
Kappa statistic                      0.8206

```

```

Mean absolute error          0.0338
Root mean squared error     0.1359
Relative absolute error     29.3237 %
Root relative squared error 56.6784 %
Total Number of Instances   3772

==== Detailed Accuracy By Class ====

      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC    ROC Area  PRC Area  Class
      0.990    0.177    0.988     0.990    0.989     0.821    0.916     0.987    negative
      0.823    0.010    0.841     0.823    0.832     0.821    0.916     0.769    sick
Weighted Avg.    0.980    0.167    0.979     0.980    0.979     0.821    0.916     0.974

==== Confusion Matrix ====

      a      b  <-- classified as
3505  36 |  a = negative
  41  190 |  b = sick

```

## Cross Validation 10 folds

```

==== Run information ===

Scheme:      weka.classifiers.meta.FilteredClassifier -F "weka.filters.supervised.attribute.Discretize -D -R first-last -precision 6" -S 1 -W
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   10-fold cross-validation

==== Classifier model (full training set) ===

FilteredClassifier using weka.classifiers.trees.J48 -C 0.25 -M 2 on data filtered through weka.filters.supervised.attribute.Discretize -D -R f

Filtered Header
@relation sick-weka.filters.supervised.attribute.Discretize-D-Rfirst-last-precision6

@attribute age_1 ('`(-inf-43.5]`','`43.5-inf)`')
@attribute age_2 ('`(-inf-69.5]`','`69.5-inf)`')
@attribute sex {F,M}
@attribute on_thyroxine {f,t}
@attribute query_on_thyroxine {f,t}
@attribute on_antithyroid_medication {f,t}
@attribute sick {f,t}
@attribute pregnant {f,t}
@attribute thyroid_surgery {f,t}
@attribute I131_treatment {f,t}
@attribute query_hypothyroid {f,t}
@attribute query_hyperthyroid {f,t}
@attribute lithium {f,t}
@attribute goitre {f,t}
@attribute tumor {f,t}
@attribute hypopituitary {f,t}
@attribute psych {f,t}
@attribute TSH_measured {t,f}

```

```

@attribute TSH {'\All'}
@attribute T3_measured {t,f}
@attribute T3_1 {'\(-inf-1.15]', '\(1.15-inf]'}
@attribute TT4_measured {t,f}
@attribute TT4_1 {'\(-inf-87.5]', '\(87.5-inf]'}
@attribute T4U_measured {t,f}
@attribute T4U_1 {'\(-inf-0.775]', '\(0.775-inf]'}
@attribute T4U_2 {'\(-inf-0.885]', '\(0.885-inf]'}
@attribute T4U_3 {'\(-inf-1.145]', '\(1.145-inf]'}
@attribute FTI_measured {t,f}
@attribute FTI {'\All'}
@attribute TBG_measured {f}
@attribute TBG {'\All'}
@attribute referral_source {SVHC,other,SVI,STMW,SVHD}
@attribute Class {negative,sick}

@data

Classifier Model
J48 pruned tree
-----
T3_1 = '\(-inf-1.15]'
| T3_measured = t
| | referral_source = SVHC: sick (9.0)
| | referral_source = other
| | | sick = f
| | | | age_2 = '\(-inf-69.5]'
| | | | on_thyroxine = f: negative (50.0/3.0)
| | | | on_thyroxine = t
| | | | | age_1 = '\(-inf-43.5)': sick (2.0)
| | | | | age_1 = '(43.5-inf)': negative (5.0/1.0)
| | | | | age_2 = '(69.5-inf)'
| | | | | T4U_measured = t: sick (19.0/6.0)
| | | | | T4U_measured = f: negative (2.0)
| | | | sick = t: sick (3.0)
| | | | referral_source = SVI
| | | | sick = f
| | | | | TT4_1 = '\(-inf-87.5]'
| | | | | T4U_measured = t
| | | | | T4U_2 = '\(-inf-0.885]'
| | | | | | query_hypothyroid = f: sick (78.0/16.0)
| | | | | | query_hypothyroid = t
| | | | | | | age_2 = '\(-inf-69.5)': sick (3.0)
| | | | | | | age_2 = '(69.5-inf)': negative (5.0/1.0)
| | | | | | | T4U_2 = '(0.885-inf)'
| | | | | | | age_1 = '\(-inf-43.5)': sick (2.0)
| | | | | | | age_1 = '(43.5-inf)'
| | | | | | | query_hypothyroid = f: negative (14.0/2.0)
| | | | | | | query_hypothyroid = t: sick (2.0)
| | | | | | | T4U_measured = f: negative (4.0)
| | | | | | TT4_1 = '(87.5-inf)': sick (90.0/10.0)
| | | | | sick = t: sick (20.0)
| | | | referral_source = STMW: sick (0.0)
| | | | referral_source = SVHD: sick (3.0)
| | T3_measured = f: negative (79.64/0.31)
T3_1 = '(1.15-inf)': negative (3381.36/24.69)

Number of Leaves : 20
Size of the tree : 36

Time taken to build model: 0.1 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3691          97.8526 %
Incorrectly Classified Instances   81           2.1474 %
Kappa statistic                   0.8113
Mean absolute error               0.0345
Root mean squared error          0.1379
Relative absolute error          29.9168 %
Root relative squared error     57.5034 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
0.989    0.186    0.988    0.989    0.989    0.811    0.913    0.986    negative
0.814    0.011    0.832    0.814    0.823    0.811    0.913    0.759    sick
Weighted Avg.  0.979    0.175    0.978    0.979    0.978    0.811    0.913    0.972

==== Confusion Matrix ====

```

```
a      b    <-- classified as
3503   38 |     a = negative
        43  188 |     b = sick
```

## Cross Validation 15 folds

```
== Run information ==
Scheme: weka.classifiers.meta.FilteredClassifier -F "weka.filters.supervised.attribute.Discretize -D -R first-last -precision 6" -S 1 -W
Relation: sick
Instances: 3772
Attributes: 30
age
sex
on_thyroxine
query_on_thyroxine
on_antithyroid_medication
sick
pregnant
thyroid_surgery
I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH
T3_measured
T3
TT4_measured
TT4
T4U_measured
T4U
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Test mode: 15-fold cross-validation
== Classifier model (full training set) ==
FilteredClassifier using weka.classifiers.trees.J48 -C 0.25 -M 2 on data filtered through weka.filters.supervised.attribute.Discretize -D -R f
Filtered Header
@relation sick-weka.filters.supervised.attribute.Discretize-D-Rfirst-last-precision6

@attribute age_1 ('\'(-inf-43.5)\',\'(43.5-inf)\')
@attribute age_2 ('\'(-inf-69.5)\',\'(69.5-inf)\')
@attribute sex {F,M}
@attribute on_thyroxine {f,t}
@attribute query_on_thyroxine {f,t}
@attribute on_antithyroid_medication {f,t}
@attribute sick {f,t}
@attribute pregnant {f,t}
@attribute thyroid_surgery {f,t}
@attribute I131_treatment {f,t}
@attribute query_hypothyroid {f,t}
@attribute query_hyperthyroid {f,t}
@attribute lithium {f,t}
@attribute goitre {f,t}
@attribute tumor {f,t}
@attribute hypopituitary {f,t}
@attribute psych {f,t}
@attribute TSH_measured {t,f}
@attribute TSH {'\All'}
@attribute T3_measured {t,f}
@attribute T3_1 {'\'(-inf-1.15)\',\'(1.15-inf)\'}
@attribute TT4_measured {t,f}
@attribute TT4_1 {'\'(-inf-87.5)\',\'(87.5-inf)\'}
@attribute T4U_measured {t,f}
@attribute T4U_1 {'\'(-inf-0.775)\',\'(0.775-inf)\'}
@attribute T4U_2 {'\'(-inf-0.885)\',\'(0.885-inf)\'}
@attribute T4U_3 {'\'(-inf-1.145)\',\'(1.145-inf)\'}
@attribute FTI_measured {t,f}
@attribute FTI {'\All'}
@attribute TBG_measured {f}
@attribute TBG {'\All'}
@attribute referral_source {SVHC,other,SVI,STMW,SVHD}
@attribute Class {negative,sick}
```

```

@data

Classifier Model
J48 pruned tree
-----

T3_1 = '(-inf-1.15)'
| T3_measured = t
| | referral_source = SVHC: sick (9.0)
| | referral_source = other
| | | sick = f
| | | | age_2 = '(-inf-69.5]'
| | | | | on_thyroxine = f: negative (50.0/3.0)
| | | | | on_thyroxine = t
| | | | | | age_1 = '(-inf-43.5]': sick (2.0)
| | | | | | age_1 = '(43.5-inf)': negative (5.0/1.0)
| | | | | | age_2 = '(69.5-inf)'
| | | | | | T4U_measured = t: sick (19.0/6.0)
| | | | | | T4U_measured = f: negative (2.0)
| | | | | sick = t: sick (3.0)
| | | | | referral_source = SVI
| | | | | sick = f
| | | | | TT4_1 = '(-inf-87.5]'
| | | | | | T4U_measured = t
| | | | | | T4U_2 = '(-inf-0.885]'
| | | | | | | query_hypothyroid = f: sick (78.0/16.0)
| | | | | | | query_hypothyroid = t
| | | | | | | | age_2 = '(-inf-69.5]': sick (3.0)
| | | | | | | | age_2 = '(69.5-inf)': negative (5.0/1.0)
| | | | | | | | T4U_2 = '(0.885-inf'
| | | | | | | | age_1 = '(-inf-43.5]': sick (2.0)
| | | | | | | | age_1 = '(43.5-inf)'
| | | | | | | | query_hypothyroid = f: negative (14.0/2.0)
| | | | | | | | query_hypothyroid = t: sick (2.0)
| | | | | | | | T4U_measured = f: negative (4.0)
| | | | | | | | TT4_1 = '(87.5-inf)': sick (90.0/10.0)
| | | | | | | sick = t: sick (20.0)
| | | | | | | referral_source = STMW: sick (0.0)
| | | | | | | referral_source = SVHD: sick (3.0)
| | T3_measured = f: negative (79.64/0.31)
T3_1 = '(1.15-inf)': negative (3381.36/24.69)

Number of Leaves : 20
Size of the tree : 36

Time taken to build model: 0.17 seconds

==== Stratified cross-validation ====
==== Summary ====

    Correctly Classified Instances      3688          97.7731 %
    Incorrectly Classified Instances   84           2.2269 %
    Kappa statistic                   0.8023
    Mean absolute error              0.0347
    Root mean squared error          0.1391
    Relative absolute error          30.1213 %
    Root relative squared error     58.0305 %
    Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.990     0.203     0.987     0.990     0.988     0.803     0.904     0.984     negative
      0.797     0.010     0.833     0.797     0.814     0.803     0.904     0.776     sick
Weighted Avg.   0.978     0.192     0.977     0.978     0.978     0.803     0.904     0.971

==== Confusion Matrix ====

      a     b  <-- classified as
3504   37 |   a = negative
      47  184 |   b = sick

```

## Percentage Split 50%

```

==== Run information ====

Scheme:      weka.classifiers.meta.FilteredClassifier -F "weka.filters.supervised.attribute.Discretize -D -R first-last -precision 6" -S 1 -W
Relation:    sick
Instances:   3772
Attributes:  30

```

```

age
sex
on_thyroxine
query_on_thyroxine
on_antithyroid_medication
sick
pregnant
thyroid_surgery
I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH
T3_measured
T3
TT4_measured
TT4
T4U_measured
T4U
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Test mode: split 50.0% train, remainder test

==== Classifier model (full training set) ====

FilteredClassifier using weka.classifiers.trees.J48 -C 0.25 -M 2 on data filtered through weka.filters.supervised.attribute.Discretize -D -R f

Filtered Header
@relation sick-weka.filters.supervised.attribute.Discretize-D-Rfirst-last-precision6

@attribute age_1 ('\'(-inf-43.5]\',',(43.5-inf)\')
@attribute age_2 ('\'(-inf-69.5]\',',(69.5-inf)\')
@attribute sex {F,M}
@attribute on_thyroxine {f,t}
@attribute query_on_thyroxine {f,t}
@attribute on_antithyroid_medication {f,t}
@attribute sick {f,t}
@attribute pregnant {f,t}
@attribute thyroid_surgery {f,t}
@attribute I131_treatment {f,t}
@attribute query_hypothyroid {f,t}
@attribute query_hyperthyroid {f,t}
@attribute lithium {f,t}
@attribute goitre {f,t}
@attribute tumor {f,t}
@attribute hypopituitary {f,t}
@attribute psych {f,t}
@attribute TSH_measured {t,f}
@attribute TSH {'\All'}
@attribute T3_measured {t,f}
@attribute T3_1 {'\'(-inf-1.15]\',',(1.15-inf)\'}
@attribute TT4_measured {t,f}
@attribute TT4_1 {'\'(-inf-87.5]\',',(87.5-inf)\'}
@attribute T4U_measured {t,f}
@attribute T4U_1 {'\'(-inf-0.775]\',',(0.775-inf)\'}
@attribute T4U_2 {'\'(-inf-0.885]\',',(0.885-inf)\'}
@attribute T4U_3 {'\'(-inf-1.145]\',',(1.145-inf)\'}
@attribute FTI_measured {t,f}
@attribute FTI {'\All'}
@attribute TBG_measured {f}
@attribute TBG {'\All'}
@attribute referral_source {SVHC,other,SVI,STMW,SVHD}
@attribute Class {negative,sick}

@data

Classifier Model
J48 pruned tree
-----
T3_1 = '(-inf-1.15)'
|   T3_measured = t
|   |   referral_source = SVHC: sick (9.0)
|   |   referral_source = other
|   |   |   sick = f
|   |   |   |   age_2 = '(-inf-69.5]'
|   |   |   |   on_thyroxine = f: negative (50.0/3.0)

```

```

|   |   |   |   |   on_thyroxine = t
|   |   |   |   |   age_1 = '(-inf-43.5]': sick (2.0)
|   |   |   |   |   age_1 = '(43.5-inf)': negative (5.0/1.0)
|   |   |   |   age_2 = '(69.5-inf)'
|   |   |   |   |   T4U_measured = t: sick (19.0/6.0)
|   |   |   |   |   T4U_measured = f: negative (2.0)
|   |   |   |   sick = t: sick (3.0)
|   referral_source = SVI
|   |   sick = f
|   |   |   TT4_1 = '(-inf-87.5]'
|   |   |   |   T4U_measured = t
|   |   |   |   |   T4U_2 = '(-inf-0.885]'
|   |   |   |   |   |   query_hypothyroid = f: sick (78.0/16.0)
|   |   |   |   |   |   query_hypothyroid = t
|   |   |   |   |   |   |   age_2 = '(-inf-69.5]': sick (3.0)
|   |   |   |   |   |   |   age_2 = '(69.5-inf)': negative (5.0/1.0)
|   |   |   |   |   |   T4U_2 = '(0.885-inf)'
|   |   |   |   |   |   |   age_1 = '(-inf-43.5]': sick (2.0)
|   |   |   |   |   |   |   age_1 = '(43.5-inf)'
|   |   |   |   |   |   |   |   query_hypothyroid = f: negative (14.0/2.0)
|   |   |   |   |   |   |   |   query_hypothyroid = t: sick (2.0)
|   |   |   |   |   |   T4U_measured = f: negative (4.0)
|   |   |   |   |   TT4_1 = '(87.5-inf)': sick (90.0/10.0)
|   |   |   |   sick = t: sick (20.0)
|   referral_source = STMW: sick (0.0)
|   |   referral_source = SVHD: sick (3.0)
|   T3_measured = f: negative (79.64/0.31)
T3_1 = '(1.15-inf)': negative (3381.36/24.69)

```

Number of Leaves : 20

Size of the tree : 36

Time taken to build model: 0.12 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

==== Summary ===

Correctly Classified Instances	1838	97.4549 %
Incorrectly Classified Instances	48	2.5451 %
Kappa statistic	0.7795	
Mean absolute error	0.0366	
Root mean squared error	0.1497	
Relative absolute error	31.6757 %	
Root relative squared error	62.0537 %	
Total Number of Instances	1886	

==== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
0.987	0.214	0.986	0.987	0.986	0.780	0.844	0.977		negative
0.786	0.013	0.800	0.786	0.793	0.780	0.844	0.701		sick
Weighted Avg.	0.975	0.201	0.974	0.975	0.974	0.780	0.844	0.960	

==== Confusion Matrix ===

a	b	<-- classified as
1746	23	a = negative
25	92	b = sick

## Percentage Split 60%

==== Run information ===

```

Scheme:      weka.classifiers.meta.FilteredClassifier -F "weka.filters.supervised.attribute.Discretize -D -R first-last -precision 6" -S 1 -W
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium

```

```

goitre
tumor
hypopituitary
psych
TSH_measured
TSH
T3_measured
T3
TT4_measured
TT4
T4U_measured
T4U
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Test mode: split 60.0% train, remainder test

==== Classifier model (full training set) ====

FilteredClassifier using weka.classifiers.trees.J48 -C 0.25 -M 2 on data filtered through weka.filters.supervised.attribute.Discretize -D -R f

Filtered Header
@relation sick-weka.filters.supervised.attribute.Discretize-D-Rfirst-last-precision6

@attribute age_1 (''(-inf-43.5)'',''(43.5-inf)'')
@attribute age_2 (''(-inf-69.5)'',''(69.5-inf)'')
@attribute sex {F,M}
@attribute on_thyroxine {f,t}
@attribute query_on_thyroxine {f,t}
@attribute on_antithyroid_medication {f,t}
@attribute sick {f,t}
@attribute pregnant {f,t}
@attribute thyroid_surgery {f,t}
@attribute I131_treatment {f,t}
@attribute query_hypothyroid {f,t}
@attribute query_hyperthyroid {f,t}
@attribute lithium {f,t}
@attribute goitre {f,t}
@attribute tumor {f,t}
@attribute hypopituitary {f,t}
@attribute psych {f,t}
@attribute TSH_measured {t,f}
@attribute TSH {'All'}
@attribute T3_measured {t,f}
@attribute T3_1 (''(-inf-1.15)'',''(1.15-inf)'')
@attribute TT4_measured {t,f}
@attribute TT4_1 (''(-inf-87.5)'',''(87.5-inf)'')
@attribute T4U_measured {t,f}
@attribute T4U_1 (''(-inf-0.775)'',''(0.775-inf)'')
@attribute T4U_2 (''(-inf-0.885)'',''(0.885-inf)'')
@attribute T4U_3 (''(-inf-1.145)'',''(1.145-inf)'')
@attribute FTI_measured {t,f}
@attribute FTI {'All'}
@attribute TBG_measured {f}
@attribute TBG {'All'}
@attribute referral_source {SVHC,other,SVI,STMW,SVHD}
@attribute Class {negative,sick}

@data

Classifier Model
J48 pruned tree
-----
T3_1 = '(-inf-1.15'
| T3_measured = t
| | referral_source = SVHC: sick (9.0)
| | referral_source = other
| | | sick = f
| | | | age_2 = '(-inf-69.5'
| | | | | on_thyroxine = f: negative (50.0/3.0)
| | | | | | age_1 = '(-inf-43.5)': sick (2.0)
| | | | | | age_1 = '(43.5-inf)': negative (5.0/1.0)
| | | | | age_2 = '(69.5-inf)'
| | | | | | T4U_measured = t: sick (19.0/6.0)
| | | | | | T4U_measured = f: negative (2.0)
| | | | | sick = t: sick (3.0)
| | | | | | referral_source = SVI
| | | | | | sick = f
| | | | | | | TT4_1 = '(-inf-87.5'
| | | | | | | T4U_measured = t
| | | | | | | | T4U_2 = '(-inf-0.885'

```

```

|   |   |   |   |   |   |   query_hypothyroid = f: sick (78.0/16.0)
|   |   |   |   |   |   |   query_hypothyroid = t
|   |   |   |   |   |   |   age_2 = '(-inf-69.5]': sick (3.0)
|   |   |   |   |   |   |   age_2 = '(69.5-inf)': negative (5.0/1.0)
|   |   |   |   |   T4U_2 = '(0.885-inf)'
|   |   |   |   |   age_1 = '(-inf-43.5]': sick (2.0)
|   |   |   |   |   age_1 = '(43.5-inf)'
|   |   |   |   |   |   query_hypothyroid = f: negative (14.0/2.0)
|   |   |   |   |   |   query_hypothyroid = t: sick (2.0)
|   |   |   |   |   T4U_measured = f: negative (4.0)
|   |   |   |   |   TT4_1 = '(87.5-inf)': sick (90.0/10.0)
|   |   |   |   |   sick = t: sick (20.0)
|   |   |   |   referral_source = STMW: sick (0.0)
|   |   |   |   referral_source = SVHD: sick (3.0)
|   |   T3_measured = f: negative (79.64/0.31)
|   |   T3_1 = '(1.15-inf)': negative (3381.36/24.69)

Number of Leaves : 20
Size of the tree : 36

Time taken to build model: 0.11 seconds
==== Evaluation on test split ====
Time taken to test model on test split: 0.02 seconds
==== Summary ====
Correctly Classified Instances      1474          97.6806 %
Incorrectly Classified Instances    35           2.3194 %
Kappa statistic                      0.7807
Mean absolute error                  0.0355
Root mean squared error              0.1427
Relative absolute error              30.9949 %
Root relative squared error         60.5594 %
Total Number of Instances           1509

==== Detailed Accuracy By Class ====


|               | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC   | ROC Area | PRC Area | Class    |
|---------------|---------|---------|-----------|--------|-----------|-------|----------|----------|----------|
| 0.991         | 0.247   | 0.985   | 0.991     | 0.988  | 0.782     | 0.906 | 0.988    | 0.988    | negative |
| 0.753         | 0.009   | 0.838   | 0.753     | 0.793  | 0.782     | 0.906 | 0.724    | 0.724    | sick     |
| Weighted Avg. | 0.977   | 0.233   | 0.976     | 0.977  | 0.976     | 0.782 | 0.906    | 0.972    |          |


==== Confusion Matrix ====


| a    | b  | <-- classified as |
|------|----|-------------------|
| 1407 | 13 | a = negative      |
| 22   | 67 | b = sick          |


```

## Percentage Split 66%

```

==== Run information ====
Scheme:      weka.classifiers.meta.FilteredClassifier -F "weka.filters.supervised.attribute.Discretize -D -R first-last -precision 6" -S 1 -W
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U

```

```

FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Test mode: split 66.0% train, remainder test

==== Classifier model (full training set) ===

FilteredClassifier using weka.classifiers.trees.J48 -C 0.25 -M 2 on data filtered through weka.filters.supervised.attribute.Discretize -D -R f

Filtered Header
@relation sick-weka.filters.supervised.attribute.Discretize-D-Rfirst-last-precision6

@attribute age_1 ('\'(-inf-43.5)\',',(43.5-inf)\')
@attribute age_2 ('\'(-inf-69.5)\',',(69.5-inf)\')
@attribute sex {F,M}
@attribute on_thyroxine {f,t}
@attribute query_on_thyroxine {f,t}
@attribute on_antithyroid_medication {f,t}
@attribute sick {f,t}
@attribute pregnant {f,t}
@attribute thyroid_surgery {f,t}
@attribute I131_treatment {f,t}
@attribute query_hypothyroid {f,t}
@attribute query_hyperthyroid {f,t}
@attribute lithium {f,t}
@attribute goitre {f,t}
@attribute tumor {f,t}
@attribute hypopituitary {f,t}
@attribute psych {f,t}
@attribute TSH_measured {t,f}
@attribute TSH {'\A1\'}
@attribute T3_measured {t,f}
@attribute T3_1 {'\'(-inf-1.15)\',',(1.15-inf)\'}
@attribute TT4_measured {t,f}
@attribute TT4_1 {'\'(-inf-87.5)\',',(87.5-inf)\'}
@attribute T4U_measured {t,f}
@attribute T4U_1 {'\'(-inf-0.775)\',',(0.775-inf)\'}
@attribute T4U_2 {'\'(-inf-0.885)\',',(0.885-inf)\'}
@attribute T4U_3 {'\'(-inf-1.145)\',',(1.145-inf)\'}
@attribute FTI_measured {t,f}
@attribute FTI {'\A1\'}
@attribute TBG_measured {f}
@attribute TBG {'\A1\'}
@attribute referral_source {SVHC,other,SVI,STMW,SVHD}
@attribute Class {negative,sick}

@data

Classifier Model
J48 pruned tree
-----
T3_1 = '(-inf-1.15'
|   T3_measured = t
|   |   referral_source = SVHC: sick (9.0)
|   |   referral_source = other
|   |   |   sick = f
|   |   |   |   age_2 = '(-inf-69.5'
|   |   |   |   |   on_thyroxine = f: negative (50.0/3.0)
|   |   |   |   |   on_thyroxine = t
|   |   |   |   |   |   age_1 = '(-inf-43.5)': sick (2.0)
|   |   |   |   |   |   age_1 = '(43.5-inf)': negative (5.0/1.0)
|   |   |   |   |   age_2 = '(69.5-inf'
|   |   |   |   |   |   T4U_measured = t: sick (19.0/6.0)
|   |   |   |   |   |   T4U_measured = f: negative (2.0)
|   |   |   |   |   sick = t: sick (3.0)
|   |   |   |   |   referral_source = SVI
|   |   |   |   |   sick = f
|   |   |   |   |   TT4_1 = '(-inf-87.5'
|   |   |   |   |   |   T4U_measured = t
|   |   |   |   |   |   T4U_2 = '(-inf-0.885'
|   |   |   |   |   |   |   query_hypothyroid = f: sick (78.0/16.0)
|   |   |   |   |   |   |   query_hypothyroid = t
|   |   |   |   |   |   |   age_2 = '(-inf-69.5)': sick (3.0)
|   |   |   |   |   |   |   age_2 = '(69.5-inf)': negative (5.0/1.0)
|   |   |   |   |   |   T4U_2 = '(0.885-inf'
|   |   |   |   |   |   |   age_1 = '(-inf-43.5)': sick (2.0)
|   |   |   |   |   |   |   age_1 = '(43.5-inf'
|   |   |   |   |   |   |   |   query_hypothyroid = f: negative (14.0/2.0)
|   |   |   |   |   |   |   |   query_hypothyroid = t: sick (2.0)
|   |   |   |   |   |   |   T4U_measured = f: negative (4.0)
|   |   |   |   |   TT4_1 = '(87.5-inf)': sick (90.0/10.0)
|   |   |   |   sick = t: sick (20.0)

```

```

|   |   referral_source = STMW: sick (0.0)
|   |   referral_source = SVHD: sick (3.0)
|   T3_measured = f: negative (79.64/0.31)
T3_1 = '(1.15-inf)': negative (3381.36/24.69)

Number of Leaves : 20

Size of the tree : 36

Time taken to build model: 0.11 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

==== Summary ===

Correctly Classified Instances      1252          97.6599 %
Incorrectly Classified Instances    30           2.3401 %
Kappa statistic                   0.7734
Mean absolute error               0.0362
Root mean squared error          0.1451
Relative absolute error          31.7024 %
Root relative squared error     61.8244 %
Total Number of Instances        1282

==== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
0.992    0.267    0.984    0.992    0.988    0.776    0.899    0.987    negative
0.733    0.008    0.846    0.733    0.786    0.776    0.899    0.711    sick
Weighted Avg.    0.977    0.252    0.976    0.977    0.976    0.776    0.899    0.971

==== Confusion Matrix ===

     a     b  <-- classified as
1197  10 |    a = negative
   20   55 |    b = sick

```

## Leave one out fold

```

==== Run information ===

Scheme:      weka.classifiers.meta.FilteredClassifier -F "weka.filters.supervised.attribute.Discretize -D -R first-last -precision 6" -S 1 -W
Relation:    sick
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   3772-fold cross-validation

==== Classifier model (full training set) ===

FilteredClassifier using weka.classifiers.trees.J48 -C 0.25 -M 2 on data filtered through weka.filters.supervised.attribute.Discretize -D -R f

```

```

Filtered Header
@relation sick-weka.filters.supervised.attribute.Discretize-D-Rfirst-last-precision6

@attribute age_1 ('`(-inf-43.5]``','`(43.5-inf)`')
@attribute age_2 ('`(-inf-69.5]``','`(69.5-inf)`')
@attribute sex {F,M}
@attribute on_thyroxine {f,t}
@attribute query_on_thyroxine {f,t}
@attribute on_antithyroid_medication {f,t}
@attribute sick {f,t}
@attribute pregnant {f,t}
@attribute thyroid_surgery {f,t}
@attribute I131_treatment {f,t}
@attribute query_hypothyroid {f,t}
@attribute query_hyperthyroid {f,t}
@attribute lithium {f,t}
@attribute goitre {f,t}
@attribute tumor {f,t}
@attribute hypopituitary {f,t}
@attribute psych {f,t}
@attribute TSH_measured {t,f}
@attribute TSH {'`All`'}
@attribute T3_measured {t,f}
@attribute T3_1 ('`(-inf-1.15]``','`(1.15-inf)`')
@attribute TT4_measured {t,f}
@attribute TT4_1 ('`(-inf-87.5]``','`(87.5-inf)`')
@attribute T4U_measured {t,f}
@attribute T4U_1 ('`(-inf-0.775]``','`(0.775-inf)`')
@attribute T4U_2 ('`(-inf-0.885]``','`(0.885-inf)`')
@attribute T4U_3 ('`(-inf-1.145]``','`(1.145-inf)`')
@attribute FTI_measured {t,f}
@attribute FTI {'`All`'}
@attribute TBG_measured {f}
@attribute TBG {'`All`'}
@attribute referral_source {SVHC,other,SVI,STMW,SVHD}
@attribute Class {negative,sick}

@data

Classifier Model
J48 pruned tree
-----
T3_1 = `(-inf-1.15'
| T3_measured = t
| | referral_source = SVHC: sick (9.0)
| | referral_source = other
| | | sick = f
| | | | age_2 = `(-inf-69.5'
| | | | | on_thyroxine = f: negative (50.0/3.0)
| | | | | on_thyroxine = t
| | | | | | age_1 = `(-inf-43.5)': sick (2.0)
| | | | | | age_1 = `(43.5-inf)': negative (5.0/1.0)
| | | | | | age_2 = `(69.5-inf'
| | | | | | T4U_measured = t: sick (19.0/6.0)
| | | | | | T4U_measured = f: negative (2.0)
| | | | | | sick = t: sick (3.0)
| | | | | | referral_source = SVI
| | | | | | sick = f
| | | | | | TT4_1 = `(-inf-87.5'
| | | | | | T4U_measured = t
| | | | | | | T4U_2 = `(-inf-0.885'
| | | | | | | | query_hypothyroid = f: sick (78.0/16.0)
| | | | | | | | query_hypothyroid = t
| | | | | | | | | age_2 = `(-inf-69.5)': sick (3.0)
| | | | | | | | | age_2 = `(69.5-inf)': negative (5.0/1.0)
| | | | | | | | | T4U_2 = `(0.885-inf'
| | | | | | | | | | age_1 = `(-inf-43.5)': sick (2.0)
| | | | | | | | | | age_1 = `(43.5-inf'
| | | | | | | | | | | query_hypothyroid = f: negative (14.0/2.0)
| | | | | | | | | | | query_hypothyroid = t: sick (2.0)
| | | | | | | | | | | T4U_measured = f: negative (4.0)
| | | | | | | | | | | TT4_1 = `(87.5-inf)': sick (90.0/10.0)
| | | | | | | | | | | sick = t: sick (20.0)
| | | | | | | | | | | referral_source = STMW: sick (0.0)
| | | | | | | | | | | referral_source = SVHD: sick (3.0)
| | | | | | | | | | | T3_measured = f: negative (79.64/0.31)
T3_1 = `(1.15-inf)': negative (3381.36/24.69)

Number of Leaves : 20

Size of the tree : 36

Time taken to build model: 0.13 seconds

```

```

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3692          97.8791 %
Incorrectly Classified Instances    80           2.1209 %
Kappa statistic                   0.814
Mean absolute error               0.0338
Root mean squared error           0.1366
Relative absolute error           29.3704 %
Root relative squared error      56.9608 %
Total Number of Instances         3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC     ROC Area   PRC Area   Class
0       0.989    0.182     0.988     0.989     0.989     0.814     0.890     0.983   negative
1       0.818    0.011     0.833     0.818     0.825     0.814     0.890     0.795   sick

Weighted Avg.      0.979    0.171     0.979     0.979     0.979     0.814     0.890     0.972

==== Confusion Matrix ====

      a      b  <-- classified as
3503   38 |  a = negative
        42  189 |  b = sick

```

### 4.3.1 Conclusion

Preprocessing method	C4.5						Leave One Out Fold	
	Cross Validation			Percentage Split				
	5 folds	10 folds	15 folds	50.00%	60.00%	66.00%		
With Original Dataset	98.4358%	98.8070%	98.6479%	98.3563%	98.6083%	98.0499%	98.5154%	
Supervised	97.9586%	97.8526%	97.7731%	97.4549%	97.6806%	97.6599%	97.8791%	

Supervised Discretizing wasn't able show any improvements in any test making it one of the worst preprocessing trick.

## 4.4 Attribute Selection

### 4.4.1 Using Filters

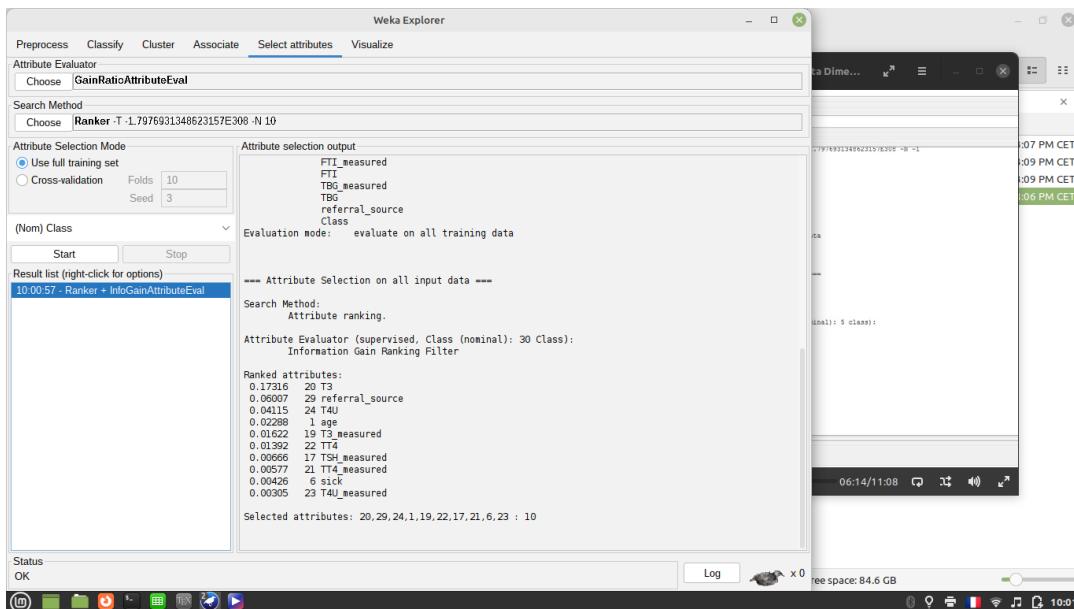


FIGURE 4.1: C4.5 Filters1

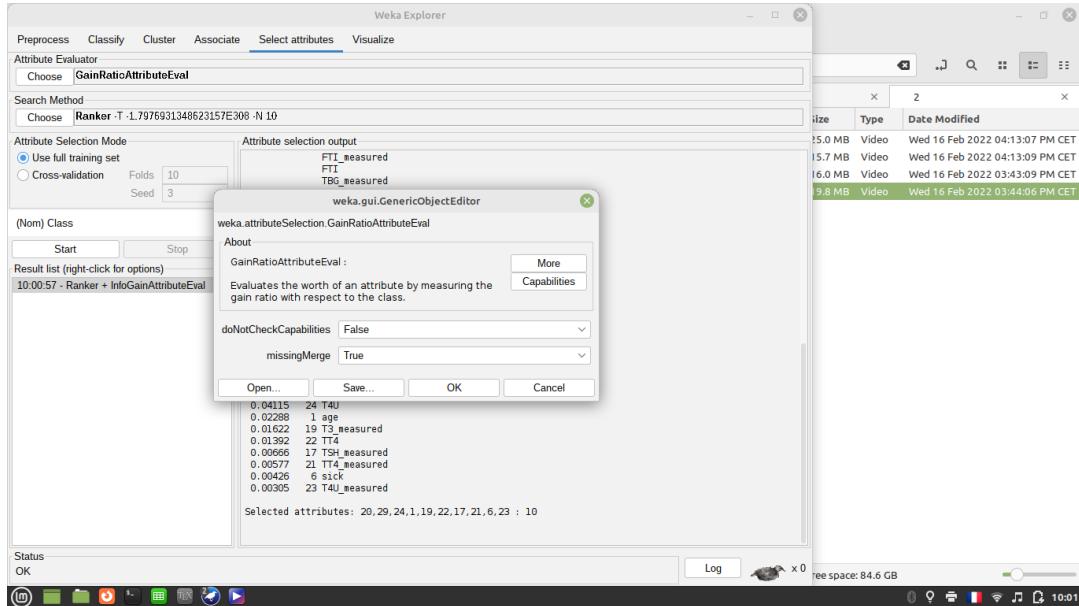


FIGURE 4.2: 4.5 Filters2

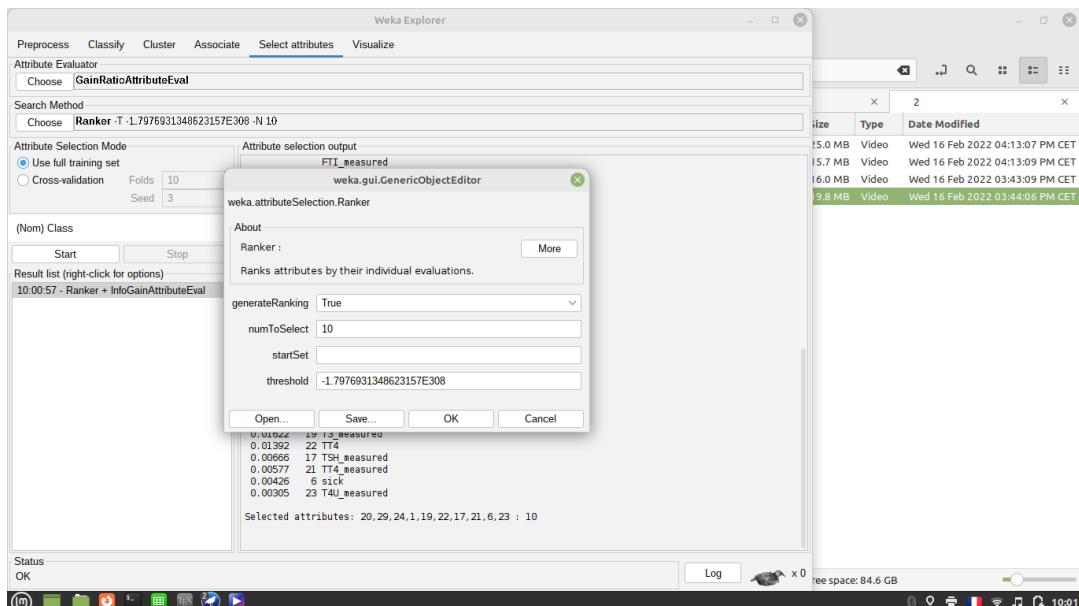


FIGURE 4.3: 4.5 Filters3

## Cross Validation 5 folds

```
==== Run information ====
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Remove-R2-5,7-16,18,25-28
Instances:   3772
Attributes:  11
              age
              sick
              TSH_measured
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              referral_source
```

```

Class
Test mode: 5-fold cross-validation

==== Classifier model (full training set) ====

J48 pruned tree
-----

T3 <= 1.1
| T3_measured = t
| | TT4 <= 53
| | | referral_source = SVHC: negative (0.0)
| | | referral_source = other: negative (30.0)
| | | referral_source = SVI
| | | | TT4 <= 31: negative (15.0)
| | | | TT4 > 31
| | | | | age <= 60
| | | | | | T4U <= 0.71: negative (6.55/2.0)
| | | | | | T4U > 0.71: sick (5.45/0.45)
| | | | | age > 60: negative (16.0/2.0)
| | | | referral_source = STMW: negative (0.0)
| | | | referral_source = SVHD: negative (0.0)
| | TT4 > 53
| | | TT4 <= 136
| | | | referral_source = SVHC: sick (9.0)
| | | | referral_source = other
| | | | | sick = f
| | | | | | age <= 77
| | | | | | | T4U <= 0.56: sick (3.28/0.28)
| | | | | | | T4U > 0.56: negative (31.72/8.0)
| | | | | | age > 77: sick (8.0)
| | | | | | sick = t: sick (3.0)
| | | | | | referral_source = SVI: sick (166.0/6.0)
| | | | | | referral_source = STMW: sick (0.0)
| | | | | | referral_source = SVHD: sick (3.0)
| | | TT4 > 136
| | | | T4U_measured = t: negative (12.0/1.0)
| | | | T4U_measured = f: sick (2.0)
| | T3_measured = f: negative (79.64/0.31)
T3 > 1.1: negative (3381.36/24.69)

Number of Leaves : 20
Size of the tree : 33

Time taken to build model: 0.2 seconds

==== Stratified cross-validation ====
==== Summary ====

  Correctly Classified Instances      3693          97.9056 %
  Incorrectly Classified Instances     79           2.0944 %
  Kappa statistic                      0.8098
  Mean absolute error                  0.0296
  Root mean squared error              0.137
  Relative absolute error              25.7276 %
  Root relative squared error         57.1424 %
  Total Number of Instances            3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC    ROC Area   PRC Area   Class
      0.992     0.216     0.986     0.992     0.989     0.811     0.916     0.987   negative
      0.784     0.008     0.862     0.784     0.821     0.811     0.916     0.780     sick
  Weighted Avg.      0.979     0.204     0.978     0.979     0.979     0.811     0.916     0.975

==== Confusion Matrix ====

      a     b  <- classified as
  3512   29 |     a = negative
      50  181 |     b = sick

```

## Cross Validation 10 folds

```

==== Run information ====

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Remove-R2-5,7-16,18,25-28
Instances:   3772
Attributes: 11
             age
             sick
             TSH_measured

```

```

T3_measured
T3
TT4_measured
TT4
T4U_measured
T4U
referral_source
Class
Test mode: 10-fold cross-validation

==== Classifier model (full training set) ===

J48 pruned tree
-----
T3 <= 1.1
| T3_measured = t
| | TT4 <= 53
| | | referral_source = SVHC: negative (0.0)
| | | referral_source = other: negative (30.0)
| | | referral_source = SVI
| | | | TT4 <= 31: negative (15.0)
| | | | TT4 > 31
| | | | | age <= 60
| | | | | | T4U <= 0.71: negative (6.55/2.0)
| | | | | | T4U > 0.71: sick (5.45/0.45)
| | | | | | age > 60: negative (16.0/2.0)
| | | | | referral_source = STMW: negative (0.0)
| | | | | referral_source = SVHD: negative (0.0)
| | | | TT4 > 53
| | | | | TT4 <= 136
| | | | | | referral_source = SVHC: sick (9.0)
| | | | | | referral_source = other
| | | | | | | sick = f
| | | | | | | age <= 77
| | | | | | | | T4U <= 0.56: sick (3.28/0.28)
| | | | | | | | T4U > 0.56: negative (31.72/8.0)
| | | | | | | | age > 77: sick (8.0)
| | | | | | | | sick = t: sick (3.0)
| | | | | | | | referral_source = SVI: sick (166.0/6.0)
| | | | | | | | referral_source = STMW: sick (0.0)
| | | | | | | | referral_source = SVHD: sick (3.0)
| | | | | TT4 > 136
| | | | | | T4U_measured = t: negative (12.0/1.0)
| | | | | | T4U_measured = f: sick (2.0)
| | T3_measured = f: negative (79.64/0.31)
T3 > 1.1: negative (3381.36/24.69)

Number of Leaves : 20

Size of the tree : 33

Time taken to build model: 0.1 seconds

==== Stratified cross-validation ====
==== Summary ===

    Correctly Classified Instances      3703           98.1707 %
    Incorrectly Classified Instances     69            1.8293 %
    Kappa statistic                      0.8317
    Mean absolute error                  0.0276
    Root mean squared error              0.1291
    Relative absolute error              23.9927 %
    Root relative squared error          53.8628 %
    Total Number of Instances            3772

==== Detailed Accuracy By Class ===

          TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
          0.994     0.208     0.987     0.994     0.990     0.834     0.911     0.985   negative
          0.792     0.006     0.897     0.792     0.841     0.834     0.911     0.783   sick
Weighted Avg.       0.982     0.195     0.981     0.982     0.981     0.834     0.911     0.973

==== Confusion Matrix ===

      a     b  <- classified as
3520  21 |   a = negative
      48  183 |   b = sick

```

## Cross Validation 15 folds

```
==== Run information ===
```

```

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Remove-R2-5,7-16,18,25-28
Instances:   3772
Attributes:  11
              age
              sick
              TSH_measured
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              referral_source
              Class
Test mode:   10-fold cross-validation

==== Classifier model (full training set) ====

J48 pruned tree
-----

T3 <= 1.1
|  T3_measured = t
|  |  TT4 <= 53
|  |  |  referral_source = SVHC: negative (0.0)
|  |  |  referral_source = other: negative (30.0)
|  |  |  referral_source = SVI
|  |  |  |  TT4 <= 31: negative (15.0)
|  |  |  |  TT4 > 31
|  |  |  |  |  age <= 60
|  |  |  |  |  |  T4U <= 0.71: negative (6.55/2.0)
|  |  |  |  |  |  T4U > 0.71: sick (5.45/0.45)
|  |  |  |  |  |  age > 60: negative (16.0/2.0)
|  |  |  |  |  referral_source = STMW: negative (0.0)
|  |  |  |  |  referral_source = SVHD: negative (0.0)
|  |  |  |  TT4 > 53
|  |  |  |  TT4 <= 136
|  |  |  |  |  referral_source = SVHC: sick (9.0)
|  |  |  |  |  referral_source = other
|  |  |  |  |  |  sick = f
|  |  |  |  |  |  age <= 77
|  |  |  |  |  |  |  T4U <= 0.56: sick (3.28/0.28)
|  |  |  |  |  |  |  T4U > 0.56: negative (31.72/8.0)
|  |  |  |  |  |  |  age > 77: sick (8.0)
|  |  |  |  |  |  |  |  sick = t: sick (3.0)
|  |  |  |  |  |  |  |  |  referral_source = SVI: sick (166.0/6.0)
|  |  |  |  |  |  |  |  |  referral_source = STMW: sick (0.0)
|  |  |  |  |  |  |  |  |  referral_source = SVHD: sick (3.0)
|  |  |  |  |  TT4 > 136
|  |  |  |  |  |  T4U_measured = t: negative (12.0/1.0)
|  |  |  |  |  |  |  T4U_measured = f: sick (2.0)
|  |  |  |  T3_measured = f: negative (79.64/0.31)
T3 > 1.1: negative (3381.36/24.69)

Number of Leaves : 20
Size of the tree : 33

Time taken to build model: 0.04 seconds

==== Stratified cross-validation ====
==== Summary ===

  Correctly Classified Instances      3703          98.1707 %
  Incorrectly Classified Instances    69           1.8293 %
  Kappa statistic                   0.8303
  Mean absolute error               0.0276
  Root mean squared error          0.1271
  Relative absolute error          24.001 %
  Root relative squared error     53.0067 %
  Total Number of Instances        3772

==== Detailed Accuracy By Class ===

          TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
  0.995    0.216    0.986    0.995    0.990    0.833    0.915    0.987    negative
  0.784    0.005    0.905    0.784    0.840    0.833    0.915    0.799    sick
Weighted Avg.  0.982    0.204    0.981    0.982    0.981    0.833    0.915    0.975

==== Confusion Matrix ====

  a      b  <- classified as
3522  19 |      a = negative
      50  181 |      b = sick

```

## Percentage Split 50%

```

==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Remove-R2-5,7-16,18,25-28
Instances:   3772
Attributes:  11
              age
              sick
              TSH_measured
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              referral_source
              Class
Test mode:   split 50.0% train, remainder test

==== Classifier model (full training set) ===

J48 pruned tree
-----

T3 <= 1.1
|  T3_measured = t
|  |  TT4 <= 53
|  |  |  referral_source = SVHC: negative (0.0)
|  |  |  referral_source = other: negative (30.0)
|  |  |  referral_source = SVI
|  |  |  |  TT4 <= 31: negative (15.0)
|  |  |  |  TT4 > 31
|  |  |  |  |  age <= 60
|  |  |  |  |  |  T4U <= 0.71: negative (6.55/2.0)
|  |  |  |  |  |  T4U > 0.71: sick (5.45/0.45)
|  |  |  |  |  age > 60: negative (16.0/2.0)
|  |  |  |  referral_source = STMW: negative (0.0)
|  |  |  |  referral_source = SVHD: negative (0.0)
|  |  |  TT4 > 53
|  |  |  TT4 <= 136
|  |  |  |  referral_source = SVHC: sick (9.0)
|  |  |  |  referral_source = other
|  |  |  |  |  sick = f
|  |  |  |  |  age <= 77
|  |  |  |  |  |  T4U <= 0.56: sick (3.28/0.28)
|  |  |  |  |  |  T4U > 0.56: negative (31.72/8.0)
|  |  |  |  |  age > 77: sick (8.0)
|  |  |  |  |  sick = t: sick (3.0)
|  |  |  |  |  referral_source = SVI: sick (166.0/6.0)
|  |  |  |  |  referral_source = STMW: sick (0.0)
|  |  |  |  |  referral_source = SVHD: sick (3.0)
|  |  |  |  TT4 > 136
|  |  |  |  |  T4U_measured = t: negative (12.0/1.0)
|  |  |  |  |  T4U_measured = f: sick (2.0)
|  |  T3_measured = f: negative (79.64/0.31)
T3 > 1.1: negative (3381.36/24.69)

Number of Leaves : 20
Size of the tree : 33

Time taken to build model: 0.06 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

==== Summary ===

  Correctly Classified Instances      1852          98.1972 %
  Incorrectly Classified Instances     34           1.8028 %
  Kappa statistic                   0.8359
  Mean absolute error                 0.0267
  Root mean squared error            0.1289
  Relative absolute error             23.1056 %
  Root relative squared error        53.4171 %
  Total Number of Instances         1886

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.994     0.205     0.987     0.994     0.990     0.838     0.919     0.989     negative
      0.795     0.006     0.903     0.795     0.845     0.838     0.919     0.786     sick
Weighted Avg.    0.982     0.193     0.981     0.982     0.981     0.838     0.919     0.976

```

```
==== Confusion Matrix ====

      a     b  <-- classified as
1759   10 |    a = negative
      24  93 |    b = sick
```

## Percentage Split 60%

```
==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Remove-R2-5,7-16,18,25-28
Instances:   3772
Attributes:  11
              age
              sick
              TSH_measured
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              referral_source
              Class
Test mode:   split 60.0% train, remainder test

==== Classifier model (full training set) ===

J48 pruned tree
-----

T3 <= 1.1
|  T3_measured = t
|  |  TT4 <= 53
|  |  |  referral_source = SVHC: negative (0.0)
|  |  |  referral_source = other: negative (30.0)
|  |  |  referral_source = SVI
|  |  |  |  TT4 <= 31: negative (15.0)
|  |  |  |  TT4 > 31
|  |  |  |  |  age <= 60
|  |  |  |  |  |  T4U <= 0.71: negative (6.55/2.0)
|  |  |  |  |  |  T4U > 0.71: sick (5.45/0.45)
|  |  |  |  |  |  age > 60: negative (16.0/2.0)
|  |  |  |  |  referral_source = STMW: negative (0.0)
|  |  |  |  |  referral_source = SVHD: negative (0.0)
|  |  |  |  TT4 > 53
|  |  |  |  |  TT4 <= 136
|  |  |  |  |  |  referral_source = SVHC: sick (9.0)
|  |  |  |  |  |  referral_source = other
|  |  |  |  |  |  |  sick = f
|  |  |  |  |  |  |  |  age <= 77
|  |  |  |  |  |  |  |  |  T4U <= 0.56: sick (3.28/0.28)
|  |  |  |  |  |  |  |  |  T4U > 0.56: negative (31.72/8.0)
|  |  |  |  |  |  |  |  |  age > 77: sick (8.0)
|  |  |  |  |  |  |  |  |  sick = t: sick (3.0)
|  |  |  |  |  |  |  |  |  referral_source = SVI: sick (166.0/6.0)
|  |  |  |  |  |  |  |  |  referral_source = STMW: sick (0.0)
|  |  |  |  |  |  |  |  |  referral_source = SVHD: sick (3.0)
|  |  |  |  |  TT4 > 136
|  |  |  |  |  |  T4U_measured = t: negative (12.0/1.0)
|  |  |  |  |  |  |  T4U_measured = f: sick (2.0)
|  |  |  |  |  T3_measured = f: negative (79.64/0.31)
T3 > 1.1: negative (3381.36/24.69)

Number of Leaves : 20
Size of the tree : 33

Time taken to build model: 0.04 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

==== Summary ===

  Correctly Classified Instances      1481          98.1445 %
  Incorrectly Classified Instances     28           1.8555 %
  Kappa statistic                      0.8235
  Mean absolute error                  0.0275
  Root mean squared error             0.1273
```

```

Relative absolute error          23.9494 %
Root relative squared error    54.0395 %
Total Number of Instances      1509

==== Detailed Accuracy By Class ====

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC     ROC Area  PRC Area  Class
0.994     0.213     0.987     0.994     0.990     0.825     0.919     0.990    negative
0.787     0.006     0.886     0.787     0.833     0.825     0.919     0.815    sick
Weighted Avg.     0.981     0.201     0.981     0.981     0.981     0.825     0.919     0.980

==== Confusion Matrix ====

      a      b  <-- classified as
1411    9 |  a = negative
    19   70 |  b = sick

```

## Percentage Split 66%

```

==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Remove-R2-5,7-16,18,25-28
Instances:   3772
Attributes:  11
              age
              sick
              TSH_measured
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              referral_source
              Class
Test mode:   split 66.0% train, remainder test

==== Classifier model (full training set) ===

J48 pruned tree
-----
T3 <= 1.1
|  T3_measured = t
|  |  TT4 <= 53
|  |  |  referral_source = SVHC: negative (0.0)
|  |  |  referral_source = other: negative (30.0)
|  |  |  referral_source = SVI
|  |  |  |  TT4 <= 31: negative (15.0)
|  |  |  |  TT4 > 31
|  |  |  |  |  age <= 60
|  |  |  |  |  |  T4U <= 0.71: negative (6.55/2.0)
|  |  |  |  |  |  T4U > 0.71: sick (5.45/0.45)
|  |  |  |  |  |  age > 60: negative (16.0/2.0)
|  |  |  |  |  |  referral_source = STMW: negative (0.0)
|  |  |  |  |  |  referral_source = SVHD: negative (0.0)
|  |  |  |  TT4 > 53
|  |  |  |  |  TT4 <= 136
|  |  |  |  |  |  referral_source = SVHC: sick (9.0)
|  |  |  |  |  |  referral_source = other
|  |  |  |  |  |  |  sick = f
|  |  |  |  |  |  |  |  age <= 77
|  |  |  |  |  |  |  |  |  T4U <= 0.56: sick (3.28/0.28)
|  |  |  |  |  |  |  |  |  T4U > 0.56: negative (31.72/8.0)
|  |  |  |  |  |  |  |  |  age > 77: sick (8.0)
|  |  |  |  |  |  |  |  |  sick = t: sick (3.0)
|  |  |  |  |  |  |  |  |  referral_source = SVI: sick (166.0/6.0)
|  |  |  |  |  |  |  |  |  referral_source = STMW: sick (0.0)
|  |  |  |  |  |  |  |  |  referral_source = SVHD: sick (3.0)
|  |  |  |  |  TT4 > 136
|  |  |  |  |  |  T4U_measured = t: negative (12.0/1.0)
|  |  |  |  |  |  T4U_measured = f: sick (2.0)
|  |  T3_measured = f: negative (79.64/0.31)
T3 > 1.1: negative (3381.36/24.69)

Number of Leaves : 20
Size of the tree : 33

Time taken to build model: 0.04 seconds

==== Evaluation on test split ====

```

```

Time taken to test model on test split: 0 seconds

==== Summary ====

Correctly Classified Instances      1256           97.9719 %
Incorrectly Classified Instances    26            2.0281 %
Kappa statistic                   0.7982
Mean absolute error               0.0288
Root mean squared error          0.1349
Relative absolute error          25.2216 %
Root relative squared error     57.4902 %
Total Number of Instances        1282

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC   ROC Area   PRC Area   Class
      0.995     0.267     0.984     0.995     0.989     0.803     0.887     0.987   negative
      0.733     0.005     0.902     0.733     0.809     0.803     0.887     0.762   sick
Weighted Avg.      0.980     0.251     0.979     0.980     0.979     0.803     0.887     0.974

==== Confusion Matrix ====

      a     b  <-- classified as
1201   6 |   a = negative
      20  55 |   b = sick

```

## Leave one out fold

```

==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Remove-R2-5,7-16,18,25-28
Instances:   3772
Attributes:  11
              age
              sick
              TSH_measured
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              referral_source
              Class
Test mode:   3772-fold cross-validation

==== Classifier model (full training set) ===

J48 pruned tree
-----
T3 <= 1.1
|  T3_measured = t
|  |  TT4 <= 53
|  |  |  referral_source = SVHC: negative (0.0)
|  |  |  referral_source = other: negative (30.0)
|  |  |  referral_source = SVI
|  |  |  |  TT4 <= 31: negative (15.0)
|  |  |  |  TT4 > 31
|  |  |  |  |  age <= 60
|  |  |  |  |  |  T4U <= 0.71: negative (6.55/2.0)
|  |  |  |  |  |  T4U > 0.71: sick (5.45/0.45)
|  |  |  |  |  age > 60: negative (16.0/2.0)
|  |  |  |  |  referral_source = STMW: negative (0.0)
|  |  |  |  |  referral_source = SVHD: negative (0.0)
|  |  |  TT4 > 53
|  |  |  |  TT4 <= 136
|  |  |  |  |  referral_source = SVHC: sick (9.0)
|  |  |  |  |  referral_source = other
|  |  |  |  |  |  sick = f
|  |  |  |  |  |  |  age <= 77
|  |  |  |  |  |  |  |  T4U <= 0.56: sick (3.28/0.28)
|  |  |  |  |  |  |  |  T4U > 0.56: negative (31.72/8.0)
|  |  |  |  |  |  |  age > 77: sick (8.0)
|  |  |  |  |  |  |  |  sick = t: sick (3.0)
|  |  |  |  |  |  |  |  referral_source = SVI: sick (166.0/6.0)
|  |  |  |  |  |  |  |  referral_source = STMW: sick (0.0)
|  |  |  |  |  |  |  |  referral_source = SVHD: sick (3.0)
|  |  |  |  TT4 > 136
|  |  |  |  |  T4U_measured = t: negative (12.0/1.0)
|  |  |  |  |  T4U_measured = f: sick (2.0)
|  |  |  |  T3_measured = f: negative (79.64/0.31)

```

```

T3 > 1.1: negative (3381.36/24.69)

Number of Leaves : 20

Size of the tree : 33

Time taken to build model: 0.05 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3702          98.1442 %
Incorrectly Classified Instances   70           1.8558 %
Kappa statistic                   0.8282
Mean absolute error               0.0273
Root mean squared error          0.1276
Relative absolute error          23.7262 %
Root relative squared error     53.2154 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC    ROC Area   PRC Area   Class
      0.994     0.216     0.986     0.994     0.990     0.830     0.902     0.988   negative
      0.784     0.006     0.900     0.784     0.838     0.830     0.902     0.815   sick
Weighted Avg.                      0.981     0.204     0.981     0.981     0.981     0.830     0.902     0.977

==== Confusion Matrix ====

      a      b  <-- classified as
3521  20 |  a = negative
      50  181 |  b = sick

```

#### 4.4.2 Using Wrapper

From the select attribute tab we are choosing **ClassifierSubsetEval** as our classifier with **PART** as the classifier option and the rest to default. Search method we are using **BestFirst** with **bidirectional** as an option.

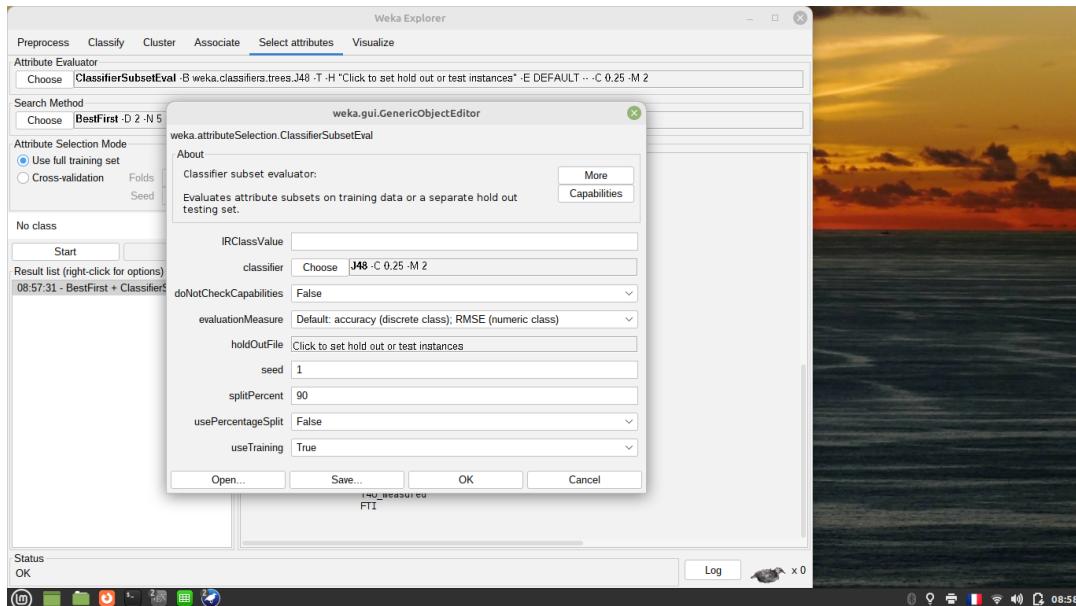


FIGURE 4.4: C4.5 Wrapper1

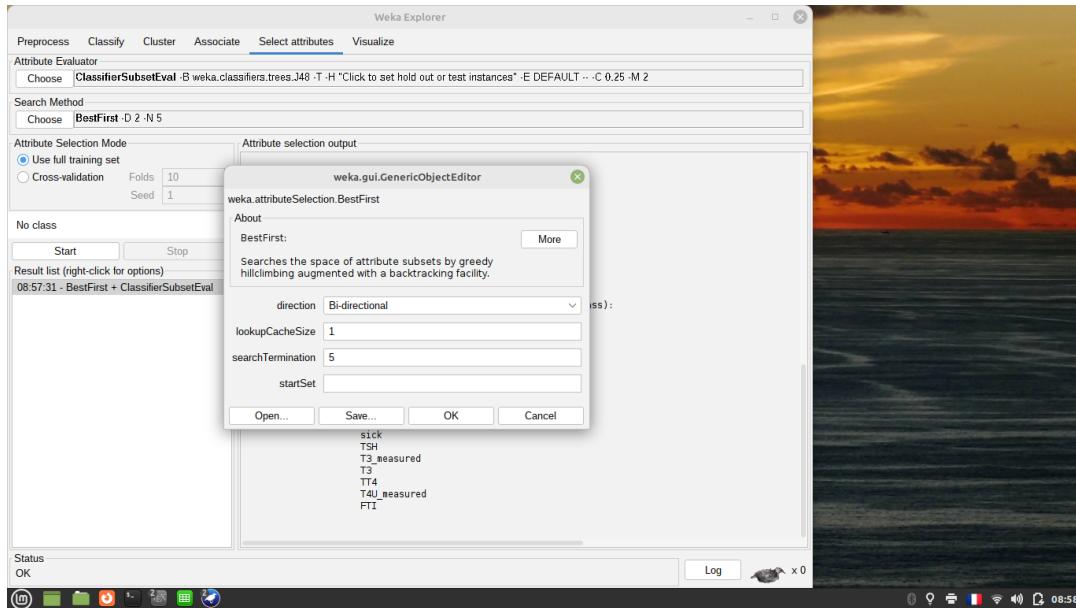


FIGURE 4.5: C4.5 Wrapper2

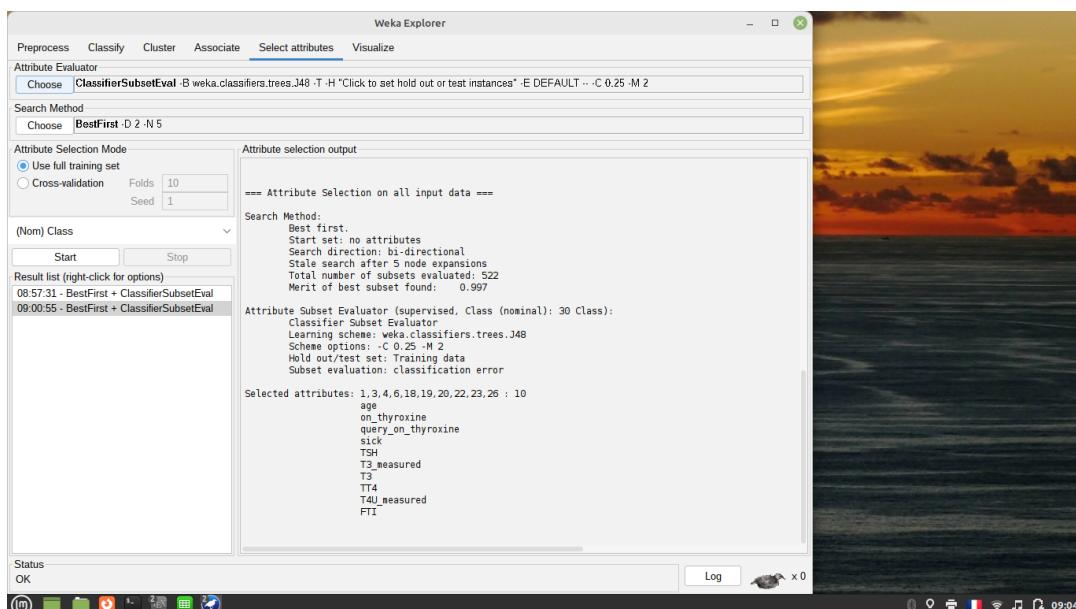


FIGURE 4.6: C4.5 Wrapper3

## Features selection

```
==== Run information ====
Evaluator: weka.attributeSelection.ClassifierSubsetEval -B weka.classifiers.trees.J48 -T -H "Click to set hold out or test instances" -E DEFAULT .. -C 0.25 -M 2
Search: weka.attributeSelection.BestFirst -D 2 -N 5
Relation: sick
Instances: 3772
Attributes: 30
age
sex
on_thyroxine
query_on_thyroxine
on_antithyroid_medication
sick
pregnant
thyroid_surgery
I131_treatment
```

```

query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH
T3_measured
T3
TT4_measured
TT4
T4U_measured
T4U
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Evaluation mode: evaluate on all training data

==== Attribute Selection on all input data ====
Search Method:
Best first.
Start set: no attributes
Search direction: bi-directional
Stale search after 5 node expansions
Total number of subsets evaluated: 522
Merit of best subset found: 0.997

Attribute Subset Evaluator (supervised, Class (nominal): 30 Class):
Classifier Subset Evaluator
Learning scheme: weka.classifiers.trees.J48
Scheme options: -C 0.25 -M 2
Hold out/test set: Training data
Subset evaluation: classification error

Selected attributes: 1,3,4,6,18,19,20,22,23,26 : 10
age
on_thyroxine
query_on_thyroxine
sick
TSH
T3_measured
T3
TT4
T4U_measured
FTI

```

## Cross Validation 5 folds

```

==== Run information ====
Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2
Relation: sick-weka.filters.unsupervised.attribute.Remove-R2,5,7-17,21,24-25,27-29
Instances: 3772
Attributes: 11
age
on_thyroxine
query_on_thyroxine
sick
TSH
T3_measured
T3
TT4
T4U_measured
FTI
Class
Test mode: 5-fold cross-validation

==== Classifier model (full training set) ====
J48 pruned tree
-----
T3 <= 1.1
|   T3_measured = t
|   |   TT4 <= 53
|   |   |   TSH <= 19

```

```

|   |   |   |   FTI <= 64: sick (6.0)
|   |   |   |   FTI > 64
|   |   |   |   |   TSH <= 8.5: negative (12.0/1.0)
|   |   |   |   |   TSH > 8.5: sick (2.0)
|   |   |   TSH > 19: negative (53.0)
|   |   TT4 > 53
|   |   |   TT4 <= 136
|   |   |   |   sick = f
|   |   |   |   |   age <= 77
|   |   |   |   |   |   query_on_thyroxine = f
|   |   |   |   |   |   |   T4U_measured = t
|   |   |   |   |   |   |   |   FTI <= 75
|   |   |   |   |   |   |   |   |   FTI <= 61: negative (4.0)
|   |   |   |   |   |   |   |   |   FTI > 61
|   |   |   |   |   |   |   |   |   |   FTI <= 69: sick (5.0)
|   |   |   |   |   |   |   |   |   |   FTI > 69: negative (4.0/1.0)
|   |   |   |   |   |   |   |   |   FTI > 75
|   |   |   |   |   |   |   |   |   |   FTI <= 156
|   |   |   |   |   |   |   |   |   |   |   age <= 44
|   |   |   |   |   |   |   |   |   |   |   |   TSH <= 3
|   |   |   |   |   |   |   |   |   |   |   |   |   TSH <= 1.7
|   |   |   |   |   |   |   |   |   |   |   |   |   |   age <= 37: sick (3.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   age > 37: negative (4.33/1.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   TSH > 1.7: negative (6.29)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   TSH > 3: sick (8.38/0.38)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   age > 44: sick (108.0/3.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   FTI > 156
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   T3 <= 0.7: sick (2.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   T3 > 0.7: negative (5.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   T4U_measured = f
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   TT4 <= 92: negative (3.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   TT4 > 92: sick (4.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   query_on_thyroxine = t
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   T3 <= 0.8: sick (2.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   T3 > 0.8: negative (2.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   age > 77: sick (38.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   sick = t: sick (25.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   TT4 > 136
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   T4U_measured = t: negative (12.0/1.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   T4U_measured = f: sick (2.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   T3_measured = f: negative (79.64/0.31)

T3 > 1.1
|   FTI <= 171
|   |   TT4 <= 56
|   |   |   T4U_measured = t
|   |   |   |   TSH <= 4.9
|   |   |   |   |   FTI <= 53: sick (8.05/1.36)
|   |   |   |   |   FTI > 53: negative (8.39/1.0)
|   |   |   |   TSH > 4.9: negative (24.66)
|   |   |   |   T4U_measured = f: negative (9.85)
|   |   |   |   TT4 > 56: negative (3187.63/1.0)
|   FTI > 171
|   |   T3 <= 2.6
|   |   |   on_thyroxine = f
|   |   |   |   T3_measured = t
|   |   |   |   |   TSH <= 0.27: sick (19.02/3.02)
|   |   |   |   |   TSH > 0.27: negative (11.76)
|   |   |   |   |   T3_measured = f: negative (8.79)
|   |   |   |   on_thyroxine = t: negative (19.64)
|   |   |   |   T3 > 2.6: negative (83.58)

Number of Leaves : 33
Size of the tree : 65

Time taken to build model: 0.09 seconds

==== Stratified cross-validation ====
==== Summary ====

    Correctly Classified Instances      3717          98.5419 %
    Incorrectly Classified Instances     55           1.4581 %
    Kappa statistic                      0.8745
    Mean absolute error                  0.0191
    Root mean squared error              0.1148
    Relative absolute error              16.5694 %
    Root relative squared error         47.8645 %
    Total Number of Instances            3772

==== Detailed Accuracy By Class ====

          TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
          0.992     0.108     0.993     0.992     0.992     0.875     0.957     0.993     negative
          0.892     0.008     0.873     0.892     0.882     0.875     0.957     0.849     sick
Weighted Avg.     0.985     0.102     0.986     0.985     0.985     0.875     0.957     0.984

```

```
==== Confusion Matrix ====
      a     b  <- classified as
3511   30 |    a = negative
      25  206 |    b = sick
```

## Cross Validation 10 folds

```
==== Run information ====
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Remove-R2,5,7-17,21,24-25,27-29
Instances:   3772
Attributes:  11
              age
              on_thyroxine
              query_on_thyroxine
              sick
              TSH
              T3_measured
              T3
              TT4
              T4U_measured
              FTI
              Class
Test mode:   10-fold cross-validation

==== Classifier model (full training set) ====
J48 pruned tree
-----
T3 <= 1.1
|  T3_measured = t
|  |  TT4 <= 53
|  |  |  TSH <= 19
|  |  |  |  FTI <= 64: sick (6.0)
|  |  |  |  FTI > 64
|  |  |  |  |  TSH <= 8.5: negative (12.0/1.0)
|  |  |  |  |  TSH > 8.5: sick (2.0)
|  |  |  |  TSH > 19: negative (53.0)
|  |  TT4 > 53
|  |  |  TT4 <= 136
|  |  |  |  sick = f
|  |  |  |  |  age <= 77
|  |  |  |  |  |  query_on_thyroxine = f
|  |  |  |  |  |  |  T4U_measured = t
|  |  |  |  |  |  |  |  FTI <= 75
|  |  |  |  |  |  |  |  |  FTI <= 61: negative (4.0)
|  |  |  |  |  |  |  |  |  FTI > 61
|  |  |  |  |  |  |  |  |  |  FTI <= 69: sick (5.0)
|  |  |  |  |  |  |  |  |  |  FTI > 69: negative (4.0/1.0)
|  |  |  |  |  |  |  |  |  FTI > 75
|  |  |  |  |  |  |  |  |  |  FTI <= 156
|  |  |  |  |  |  |  |  |  |  |  age <= 44
|  |  |  |  |  |  |  |  |  |  |  |  TSH <= 3
|  |  |  |  |  |  |  |  |  |  |  |  |  TSH <= 1.7
|  |  |  |  |  |  |  |  |  |  |  |  |  |  age <= 37: sick (3.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  age > 37: negative (4.33/1.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  TSH > 1.7: negative (6.29)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  TSH > 3: sick (8.38/0.38)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  age > 44: sick (108.0/3.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  FTI > 156
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  T3 <= 0.7: sick (2.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  T3 > 0.7: negative (5.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  T4U_measured = f
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  TT4 <= 92: negative (3.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  TT4 > 92: sick (4.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  query_on_thyroxine = t
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  T3 <= 0.8: sick (2.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  T3 > 0.8: negative (2.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  age > 77: sick (38.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  sick = t: sick (25.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  TT4 > 136
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  T4U_measured = t: negative (12.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  T4U_measured = f: sick (2.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  T3_measured = f: negative (79.64/0.31)
T3 > 1.1
|  FTI <= 171
|  |  TT4 <= 56
|  |  |  T4U_measured = t
|  |  |  |  TSH <= 4.9
|  |  |  |  |  FTI <= 53: sick (8.05/1.36)
|  |  |  |  |  FTI > 53: negative (8.39/1.0)
```

```

|   |   |   |   TSH > 4.9: negative (24.66)
|   |   |   |   T4U_measured = f: negative (9.85)
|   |   |   TT4 > 56: negative (3187.63/1.0)
|   FTI > 171
|   |   T3 <= 2.6
|   |   |   on_thyroxine = f
|   |   |   |   T3_measured = t
|   |   |   |   |   TSH <= 0.27: sick (19.02/3.02)
|   |   |   |   |   TSH > 0.27: negative (11.76)
|   |   |   |   T3_measured = f: negative (8.79)
|   |   |   |   on_thyroxine = t: negative (19.64)
|   |   |   T3 > 2.6: negative (83.58)

Number of Leaves : 33

Size of the tree : 65

Time taken to build model: 0.07 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3715          98.4889 %
Incorrectly Classified Instances    57           1.5111 %
Kappa statistic                   0.8709
Mean absolute error               0.0189
Root mean squared error          0.1171
Relative absolute error          16.37 %
Root relative squared error     48.8315 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC   ROC Area   PRC Area   Class
      0.991     0.104     0.993     0.991     0.992     0.871     0.950     0.992   negative
      0.896     0.009     0.863     0.896     0.879     0.871     0.950     0.827   sick
Weighted Avg.      0.985     0.098     0.985     0.985     0.985     0.871     0.950     0.982

==== Confusion Matrix ====

      a     b  <- classified as
3508   33 |   a = negative
      24  207 |   b = sick

```

## Cross Validation 15 folds

```

==== Run information ====

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Remove-R2,5,7-17,21,24-25,27-29
Instances:   3772
Attributes:  11
              age
              on_thyroxine
              query_on_thyroxine
              sick
              TSH
              T3_measured
              T3
              TT4
              T4U_measured
              FTI
              Class
Test mode:   15-fold cross-validation

==== Classifier model (full training set) ====

J48 pruned tree
-----
T3 <= 1.1
|   T3_measured = t
|   |   TT4 <= 53
|   |   |   TSH <= 19
|   |   |   |   FTI <= 64: sick (6.0)
|   |   |   |   FTI > 64
|   |   |   |   |   TSH <= 8.5: negative (12.0/1.0)
|   |   |   |   |   TSH > 8.5: sick (2.0)
|   |   |   |   TSH > 19: negative (53.0)
|   |   TT4 > 53
|   |   |   TT4 <= 136
|   |   |   |   sick = f
|   |   |   |   |   age <= 77

```

```

|   |   |   |   |   query_on_thyroxine = f
|   |   |   |   |   |   T4U_measured = t
|   |   |   |   |   |   |   FTI <= 75
|   |   |   |   |   |   |   |   FTI <= 61: negative (4.0)
|   |   |   |   |   |   |   |   FTI > 61
|   |   |   |   |   |   |   |   |   FTI <= 69: sick (5.0)
|   |   |   |   |   |   |   |   |   FTI > 69: negative (4.0/1.0)
|   |   |   |   |   |   |   |   FTI > 75
|   |   |   |   |   |   |   |   |   FTI <= 156
|   |   |   |   |   |   |   |   |   age <= 44
|   |   |   |   |   |   |   |   |   |   TSH <= 3
|   |   |   |   |   |   |   |   |   |   |   TSH <= 1.7
|   |   |   |   |   |   |   |   |   |   |   |   age <= 37: sick (3.0)
|   |   |   |   |   |   |   |   |   |   |   |   age > 37: negative (4.33/1.0)
|   |   |   |   |   |   |   |   |   |   |   TSH > 1.7: negative (6.29)
|   |   |   |   |   |   |   |   |   |   TSH > 3: sick (8.38/0.38)
|   |   |   |   |   |   |   |   |   |   |   age > 44: sick (108.0/3.0)
|   |   |   |   |   |   |   |   |   |   FTI > 156
|   |   |   |   |   |   |   |   |   |   |   T3 <= 0.7: sick (2.0)
|   |   |   |   |   |   |   |   |   |   |   T3 > 0.7: negative (5.0)
|   |   |   |   |   |   |   |   |   |   T4U_measured = f
|   |   |   |   |   |   |   |   |   |   TT4 <= 92: negative (3.0)
|   |   |   |   |   |   |   |   |   |   TT4 > 92: sick (4.0)
|   |   |   |   |   |   |   |   |   query_on_thyroxine = t
|   |   |   |   |   |   |   |   |   |   T3 <= 0.8: sick (2.0)
|   |   |   |   |   |   |   |   |   |   T3 > 0.8: negative (2.0)
|   |   |   |   |   |   |   |   |   |   |   age > 77: sick (38.0)
|   |   |   |   |   |   |   |   |   |   |   sick = t: sick (25.0)
|   |   |   |   |   |   |   |   |   |   TT4 > 136
|   |   |   |   |   |   |   |   |   |   |   T4U_measured = t: negative (12.0/1.0)
|   |   |   |   |   |   |   |   |   |   |   T4U_measured = f: sick (2.0)
|   |   |   |   |   |   |   |   |   |   T3_measured = f: negative (79.64/0.31)
T3 > 1.1
|   |   |   |   |   |   |   |   |   |   FTI <= 171
|   |   |   |   |   |   |   |   |   |   TT4 <= 56
|   |   |   |   |   |   |   |   |   |   T4U_measured = t
|   |   |   |   |   |   |   |   |   |   TSH <= 4.9
|   |   |   |   |   |   |   |   |   |   |   FTI <= 53: sick (8.05/1.36)
|   |   |   |   |   |   |   |   |   |   |   FTI > 53: negative (8.39/1.0)
|   |   |   |   |   |   |   |   |   |   TSH > 4.9: negative (24.66)
|   |   |   |   |   |   |   |   |   |   T4U_measured = f: negative (9.85)
|   |   |   |   |   |   |   |   |   |   TT4 > 56: negative (3187.63/1.0)
|   |   |   |   |   |   |   |   |   FTI > 171
|   |   |   |   |   |   |   |   |   |   T3 <= 2.6
|   |   |   |   |   |   |   |   |   |   on_thyroxine = f
|   |   |   |   |   |   |   |   |   |   T3_measured = t
|   |   |   |   |   |   |   |   |   |   |   TSH <= 0.27: sick (19.02/3.02)
|   |   |   |   |   |   |   |   |   |   |   TSH > 0.27: negative (11.76)
|   |   |   |   |   |   |   |   |   |   T3_measured = f: negative (8.79)
|   |   |   |   |   |   |   |   |   |   on_thyroxine = t: negative (19.64)
|   |   |   |   |   |   |   |   |   |   T3 > 2.6: negative (83.58)

Number of Leaves : 33
Size of the tree : 65

Time taken to build model: 0.12 seconds

==== Stratified cross-validation ====
==== Summary ====

    Correctly Classified Instances      3720          98.6214 %
    Incorrectly Classified Instances     52           1.3786 %
    Kappa statistic                      0.8801
    Mean absolute error                  0.0183
    Root mean squared error              0.112
    Relative absolute error              15.8677 %
    Root relative squared error         46.7259 %
    Total Number of Instances            3772

==== Detailed Accuracy By Class ====

          TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
          0.993     0.113     0.993     0.993     0.993     0.880     0.954     0.992     negative
          0.887     0.007     0.887     0.887     0.887     0.880     0.954     0.848     sick
Weighted Avg.   0.986     0.106     0.986     0.986     0.986     0.880     0.954     0.983

==== Confusion Matrix ====

      a      b  <-- classified as
3515  26 |  a = negative
      26  205 |  b = sick

```

## Percentage Split 50%

```

==== Run information ====

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Remove-R2,5,7-17,21,24-25,27-29
Instances:   3772
Attributes:  11
              age
              on_thyroxine
              query_on_thyroxine
              sick
              TSH
              T3_measured
              T3
              TT4
              T4U_measured
              FTI
              Class
Test mode:   split 50.0% train, remainder test

==== Classifier model (full training set) ====

J48 pruned tree
-----
T3 <= 1.1
|  T3_measured = t
|  |  TT4 <= 53
|  |  |  TSH <= 19
|  |  |  |  FTI <= 64: sick (6.0)
|  |  |  |  FTI > 64
|  |  |  |  |  TSH <= 8.5: negative (12.0/1.0)
|  |  |  |  |  TSH > 8.5: sick (2.0)
|  |  |  |  |  TSH > 19: negative (53.0)
|  |  |  TT4 > 53
|  |  |  |  TT4 <= 136
|  |  |  |  |  sick = f
|  |  |  |  |  |  age <= 77
|  |  |  |  |  |  |  query_on_thyroxine = f
|  |  |  |  |  |  |  |  T4U_measured = t
|  |  |  |  |  |  |  |  |  FTI <= 75
|  |  |  |  |  |  |  |  |  |  FTI <= 61: negative (4.0)
|  |  |  |  |  |  |  |  |  |  FTI > 61
|  |  |  |  |  |  |  |  |  |  |  FTI <= 69: sick (5.0)
|  |  |  |  |  |  |  |  |  |  |  FTI > 69: negative (4.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  FTI > 75
|  |  |  |  |  |  |  |  |  |  |  |  FTI <= 156
|  |  |  |  |  |  |  |  |  |  |  |  |  age <= 44
|  |  |  |  |  |  |  |  |  |  |  |  |  |  TSH <= 3
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  TSH <= 1.7
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  age <= 37: sick (3.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  age > 37: negative (4.33/1.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  TSH > 1.7: negative (6.29)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  TSH > 3: sick (8.38/0.38)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  age > 44: sick (108.0/3.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  FTI > 156
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  T3 <= 0.7: sick (2.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  T3 > 0.7: negative (5.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  T4U_measured = f
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  TT4 <= 92: negative (3.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  TT4 > 92: sick (4.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  query_on_thyroxine = t
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  T3 <= 0.8: sick (2.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  T3 > 0.8: negative (2.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  age > 77: sick (38.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  sick = t: sick (25.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  TT4 > 136
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  T4U_measured = t: negative (12.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  T4U_measured = f: sick (2.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  T3_measured = f: negative (79.64/0.31)

T3 > 1.1
|  FTI <= 171
|  |  TT4 <= 56
|  |  |  T4U_measured = t
|  |  |  |  TSH <= 4.9
|  |  |  |  |  FTI <= 53: sick (8.05/1.36)
|  |  |  |  |  |  FTI > 53: negative (8.39/1.0)
|  |  |  |  |  TSH > 4.9: negative (24.66)
|  |  |  |  |  T4U_measured = f: negative (9.85)
|  |  |  |  |  TT4 > 56: negative (3187.63/1.0)
|  |  |  |  FTI > 171
|  |  |  |  T3 <= 2.6
|  |  |  |  |  on_thyroxine = f
|  |  |  |  |  |  T3_measured = t
|  |  |  |  |  |  |  TSH <= 0.27: sick (19.02/3.02)
|  |  |  |  |  |  |  |  TSH > 0.27: negative (11.76)

```

```

|   |   |   |   T3_measured = f: negative (8.79)
|   |   |   on_thyroxine = t: negative (19.64)
|   |   T3 > 2.6: negative (83.58)

Number of Leaves : 33

Size of the tree : 65

Time taken to build model: 0.06 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

==== Summary ===

Correctly Classified Instances      1856           98.4093 %
Incorrectly Classified Instances    30            1.5907 %
Kappa statistic                      0.8675
Mean absolute error                  0.0227
Root mean squared error              0.1204
Relative absolute error              19.6372 %
Root relative squared error         49.9322 %
Total Number of Instances          1886

==== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC     ROC Area  PRC Area  Class
      0.989    0.094    0.994    0.989    0.992    0.868    0.954    0.993    negative
      0.906    0.011    0.848    0.906    0.876    0.868    0.954    0.802    sick
Weighted Avg.  0.984    0.089    0.985    0.984    0.984    0.868    0.954    0.981

==== Confusion Matrix ===

      a      b  <-- classified as
1750    19 |      a = negative
      11   106 |      b = sick

```

## Percentage Split 60%

```

==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Remove-R2,5,7-17,21,24-25,27-29
Instances:   3772
Attributes:  11
              age
              on_thyroxine
              query_on_thyroxine
              sick
              TSH
              T3_measured
              T3
              TT4
              T4U_measured
              FTI
              Class
Test mode:   split 60.0% train, remainder test

==== Classifier model (full training set) ===

J48 pruned tree
-----

T3 <= 1.1
|   T3_measured = t
|   |   TT4 <= 53
|   |   |   TSH <= 19
|   |   |   |   FTI <= 64: sick (6.0)
|   |   |   |   FTI > 64
|   |   |   |   |   TSH <= 8.5: negative (12.0/1.0)
|   |   |   |   |   TSH > 8.5: sick (2.0)
|   |   |   |   TSH > 19: negative (53.0)
|   |   |   TT4 > 53
|   |   |   |   TT4 <= 136
|   |   |   |   |   sick = f
|   |   |   |   |   |   age <= 77
|   |   |   |   |   |   |   query_on_thyroxine = f
|   |   |   |   |   |   |   T4U_measured = t
|   |   |   |   |   |   |   |   FTI <= 75
|   |   |   |   |   |   |   |   |   FTI <= 61: negative (4.0)
|   |   |   |   |   |   |   |   |   FTI > 61
|   |   |   |   |   |   |   |   |   |   FTI <= 69: sick (5.0)

```

```

|   |   |   |   |   |   |   |   FTI > 69: negative (4.0/1.0)
|   |   |   |   |   |   |   |   FTI > 75
|   |   |   |   |   |   |   |   FTI <= 156
|   |   |   |   |   |   |   |   age <= 44
|   |   |   |   |   |   |   |   TSH <= 3
|   |   |   |   |   |   |   |   TSH <= 1.7
|   |   |   |   |   |   |   |   age <= 37: sick (3.0)
|   |   |   |   |   |   |   |   age > 37: negative (4.33/1.0)
|   |   |   |   |   |   |   |   TSH > 1.7: negative (6.29)
|   |   |   |   |   |   |   |   TSH > 3: sick (8.38/0.38)
|   |   |   |   |   |   |   |   age > 44: sick (108.0/3.0)
|   |   |   |   |   |   |   |   FTI > 156
|   |   |   |   |   |   |   |   T3 <= 0.7: sick (2.0)
|   |   |   |   |   |   |   |   T3 > 0.7: negative (5.0)
|   |   |   |   |   T4U_measured = f
|   |   |   |   |   TT4 <= 92: negative (3.0)
|   |   |   |   |   TT4 > 92: sick (4.0)
|   |   |   |   query_on_thyroxine = t
|   |   |   |   T3 <= 0.8: sick (2.0)
|   |   |   |   T3 > 0.8: negative (2.0)
|   |   |   |   age > 77: sick (38.0)
|   |   |   |   sick = t: sick (25.0)
|   |   |   TT4 > 136
|   |   |   T4U_measured = t: negative (12.0/1.0)
|   |   |   T4U_measured = f: sick (2.0)
|   |   T3_measured = f: negative (79.64/0.31)

T3 > 1.1
|   FTI <= 171
|   |   TT4 <= 56
|   |   |   T4U_measured = t
|   |   |   |   TSH <= 4.9
|   |   |   |   FTI <= 53: sick (8.05/1.36)
|   |   |   |   FTI > 53: negative (8.39/1.0)
|   |   |   |   TSH > 4.9: negative (24.66)
|   |   |   T4U_measured = f: negative (9.85)
|   |   TT4 > 56: negative (3187.63/1.0)
|   FTI > 171
|   |   T3 <= 2.6
|   |   |   on_thyroxine = f
|   |   |   |   T3_measured = t
|   |   |   |   TSH <= 0.27: sick (19.02/3.02)
|   |   |   |   TSH > 0.27: negative (11.76)
|   |   |   |   T3_measured = f: negative (8.79)
|   |   |   on_thyroxine = t: negative (19.64)
|   |   T3 > 2.6: negative (83.58)

Number of Leaves : 33

Size of the tree : 65

Time taken to build model: 0.06 seconds

== Evaluation on test split ==

Time taken to test model on test split: 0 seconds

== Summary ==

Correctly Classified Instances      1487      98.5421 %
Incorrectly Classified Instances     22       1.4579 %
Kappa statistic                      0.8714
Mean absolute error                  0.0214
Root mean squared error              0.1162
Relative absolute error              18.639 %
Root relative squared error         49.3272 %
Total Number of Instances           1509

== Detailed Accuracy By Class ==

          TP Rate  FP Rate  Precision  Recall   F-Measure  MCC    ROC Area  PRC Area  Class
0.991      0.101     0.994     0.991     0.992     0.872     0.955     0.994  negative
0.899      0.009     0.860     0.899     0.879     0.872     0.955     0.798   sick
Weighted Avg.    0.985     0.096     0.986     0.985     0.986     0.872     0.955     0.983

== Confusion Matrix ==

a      b      <- classified as
1407   13 |      a = negative
      9   80 |      b = sick
```

## Percentage Split 66%

### ==== Run information ====

```

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Remove-R2,5,7-17,21,24-25,27-29
Instances:   3772
Attributes:  11
             age
             on_thyroxine
             query_on_thyroxine
             sick
             TSH
             T3_measured
             T3
             TT4
             T4U_measured
             FTI
             Class
Test mode:   split 66.0% train, remainder test

==== Classifier model (full training set) ====

J48 pruned tree
-----
T3 <= 1.1
|  T3_measured = t
|  |  TT4 <= 53
|  |  |  TSH <= 19
|  |  |  |  FTI <= 64: sick (6.0)
|  |  |  |  FTI > 64
|  |  |  |  |  TSH <= 8.5: negative (12.0/1.0)
|  |  |  |  |  |  TSH > 8.5: sick (2.0)
|  |  |  |  TSH > 19: negative (53.0)
|  |  TT4 > 53
|  |  |  TT4 <= 136
|  |  |  |  sick = f
|  |  |  |  |  age <= 77
|  |  |  |  |  |  query_on_thyroxine = f
|  |  |  |  |  |  |  T4U_measured = t
|  |  |  |  |  |  |  |  FTI <= 75
|  |  |  |  |  |  |  |  |  FTI <= 61: negative (4.0)
|  |  |  |  |  |  |  |  |  FTI > 61
|  |  |  |  |  |  |  |  |  |  FTI <= 69: sick (5.0)
|  |  |  |  |  |  |  |  |  |  FTI > 69: negative (4.0/1.0)
|  |  |  |  |  |  |  |  |  FTI > 75
|  |  |  |  |  |  |  |  |  |  FTI <= 156
|  |  |  |  |  |  |  |  |  |  |  age <= 44
|  |  |  |  |  |  |  |  |  |  |  |  TSH <= 3
|  |  |  |  |  |  |  |  |  |  |  |  |  TSH <= 1.7
|  |  |  |  |  |  |  |  |  |  |  |  |  |  age <= 37: sick (3.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  age > 37: negative (4.33/1.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  TSH > 1.7: negative (6.29)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  TSH > 3: sick (8.38/0.38)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  age > 44: sick (108.0/3.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  FTI > 156
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  T3 <= 0.7: sick (2.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  T3 > 0.7: negative (5.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  T4U_measured = f
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  TT4 <= 92: negative (3.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  TT4 > 92: sick (4.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  query_on_thyroxine = t
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  T3 <= 0.8: sick (2.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  T3 > 0.8: negative (2.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  age > 77: sick (38.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  sick = t: sick (25.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  TT4 > 136
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  T4U_measured = t: negative (12.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  T4U_measured = f: sick (2.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  T3_measured = f: negative (79.64/0.31)

T3 > 1.1
|  FTI <= 171
|  |  TT4 <= 56
|  |  |  T4U_measured = t
|  |  |  |  TSH <= 4.9
|  |  |  |  |  FTI <= 53: sick (8.05/1.36)
|  |  |  |  |  FTI > 53: negative (8.39/1.0)
|  |  |  |  TSH > 4.9: negative (24.66)
|  |  |  |  T4U_measured = f: negative (9.85)
|  |  |  TT4 > 56: negative (3187.63/1.0)
|  |  FTI > 171
|  |  |  T3 <= 2.6
|  |  |  |  on_thyroxine = f
|  |  |  |  |  T3_measured = t
|  |  |  |  |  |  TSH <= 0.27: sick (19.02/3.02)
|  |  |  |  |  |  TSH > 0.27: negative (11.76)
|  |  |  |  |  |  T3_measured = f: negative (8.79)
|  |  |  |  on_thyroxine = t: negative (19.64)
|  |  |  |  T3 > 2.6: negative (83.58)

```

```

Number of Leaves : 33
Size of the tree : 65

Time taken to build model: 0.06 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0.01 seconds

==== Summary ===

Correctly Classified Instances      1270          99.064 %
Incorrectly Classified Instances   12           0.936 %
Kappa statistic                   0.9161
Mean absolute error               0.0185
Root mean squared error          0.0944
Relative absolute error          16.1984 %
Root relative squared error     40.2108 %
Total Number of Instances        1282

==== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC    ROC Area  PRC Area  Class
0  0.994    0.067    0.996    0.994    0.995    0.916  0.971    0.997  negative
0  0.933    0.006    0.909    0.933    0.921    0.916  0.971    0.864  sick
Weighted Avg.  0.991    0.063    0.991    0.991    0.991    0.916  0.971    0.989

==== Confusion Matrix ===

     a      b  <-- classified as
1200    7 |      a = negative
      5    70 |      b = sick

```

## Leave one out fold

```

==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.Remove-R2,5,7-17,21,24-25,27-29
Instances:   3772
Attributes:  11
              age
              on_thyroxine
              query_on_thyroxine
              sick
              TSH
              T3_measured
              T3
              TT4
              T4U_measured
              FTI
              Class
Test mode:   3772-fold cross-validation

==== Classifier model (full training set) ===

J48 pruned tree
-----
T3 <= 1.1
|  T3_measured = t
|  |  TT4 <= 53
|  |  |  TSH <= 19
|  |  |  |  FTI <= 64: sick (6.0)
|  |  |  |  FTI > 64
|  |  |  |  |  TSH <= 8.5: negative (12.0/1.0)
|  |  |  |  |  TSH > 8.5: sick (2.0)
|  |  |  |  TSH > 19: negative (53.0)
|  |  TT4 > 53
|  |  |  TT4 <= 136
|  |  |  |  sick = f
|  |  |  |  age <= 77
|  |  |  |  |  query_on_thyroxine = f
|  |  |  |  |  |  T4U_measured = t
|  |  |  |  |  |  |  FTI <= 75
|  |  |  |  |  |  |  |  FTI <= 61: negative (4.0)
|  |  |  |  |  |  |  |  FTI > 61
|  |  |  |  |  |  |  |  |  FTI <= 69: sick (5.0)
|  |  |  |  |  |  |  |  |  FTI > 69: negative (4.0/1.0)
|  |  |  |  |  |  |  |  |  FTI > 75
|  |  |  |  |  |  |  |  |  FTI <= 156

```

```

|   |   |   |   |   |   |   |   |   age <= 44
|   |   |   |   |   |   |   |   |   TSH <= 3
|   |   |   |   |   |   |   |   |   TSH <= 1.7
|   |   |   |   |   |   |   |   |   age <= 37: sick (3.0)
|   |   |   |   |   |   |   |   |   age > 37: negative (4.33/1.0)
|   |   |   |   |   |   |   |   |   TSH > 1.7: negative (6.29)
|   |   |   |   |   |   |   |   |   TSH > 3: sick (8.38/0.38)
|   |   |   |   |   |   |   |   |   age > 44: sick (108.0/3.0)
|   |   |   |   |   FTI > 156
|   |   |   |   |   |   T3 <= 0.7: sick (2.0)
|   |   |   |   |   |   T3 > 0.7: negative (5.0)
|   |   |   |   |   T4U_measured = f
|   |   |   |   |   |   TT4 <= 92: negative (3.0)
|   |   |   |   |   |   TT4 > 92: sick (4.0)
|   |   |   |   |   query_on_thyroxine = t
|   |   |   |   |   |   T3 <= 0.8: sick (2.0)
|   |   |   |   |   |   T3 > 0.8: negative (2.0)
|   |   |   |   |   age > 77: sick (38.0)
|   |   |   |   |   sick = t: sick (25.0)
|   |   |   |   TT4 > 136
|   |   |   |   T4U_measured = t: negative (12.0/1.0)
|   |   |   |   T4U_measured = f: sick (2.0)
|   |   |   T3_measured = f: negative (79.64/0.31)

T3 > 1.1
|   FTI <= 171
|   |   TT4 <= 56
|   |   |   T4U_measured = t
|   |   |   |   TSH <= 4.9
|   |   |   |   |   FTI <= 53: sick (8.05/1.36)
|   |   |   |   |   FTI > 53: negative (8.39/1.0)
|   |   |   |   |   TSH > 4.9: negative (24.66)
|   |   |   |   T4U_measured = f: negative (9.85)
|   |   |   TT4 > 56: negative (3187.63/1.0)
|   FTI > 171
|   |   T3 <= 2.6
|   |   |   on_thyroxine = f
|   |   |   |   T3_measured = t
|   |   |   |   |   TSH <= 0.27: sick (19.02/3.02)
|   |   |   |   |   TSH > 0.27: negative (11.76)
|   |   |   |   |   T3_measured = f: negative (8.79)
|   |   |   |   on_thyroxine = t: negative (19.64)
|   |   |   T3 > 2.6: negative (83.58)

Number of Leaves : 33
Size of the tree : 65

Time taken to build model: 0.07 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3717          98.5419 %
Incorrectly Classified Instances     55           1.4581 %
Kappa statistic                      0.8729
Mean absolute error                  0.0173
Root mean squared error              0.1158
Relative absolute error              15.0439 %
Root relative squared error         48.2685 %
Total Number of Instances            3772

==== Detailed Accuracy By Class ====

          TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
          0.992    0.121    0.992    0.992    0.992    0.873    0.934    0.988    negative
          0.879    0.008    0.883    0.879    0.881    0.873    0.934    0.865    sick
Weighted Avg.    0.985    0.114    0.985    0.985    0.985    0.873    0.934    0.981

==== Confusion Matrix ====

      a      b  <- classified as
3514  27 |    a = negative
      28  203 |    b = sick

```

### 4.4.3 Conclusion

Removing attributes with the wrapper method was able to show improvements in performances in 4 tests out 7 which make it a useful preprocessing technique, however with the filter method we couldn't see any improvement.

Preprocessing method	Cross Validation			Percentage Split			Leave One Out Fold
	5 folds	10 folds	15 folds	50.00%	60.00%	66.00%	
With Original Dataset	98.4358%	98.8070%	98.6479%	98.3563%	98.6083%	98.0499%	98.5154%
With Filter	97.9056%	98.1707%	98.1707%	98.1972%	98.1445%	97.9719%	98.1442%
With Wrapper	98.5419%	98.4889%	98.6214%	98.4093%	98.5421%	99.0640%	98.5419%

## 4.5 Data Cleaning

### 4.5.1 Missing Values

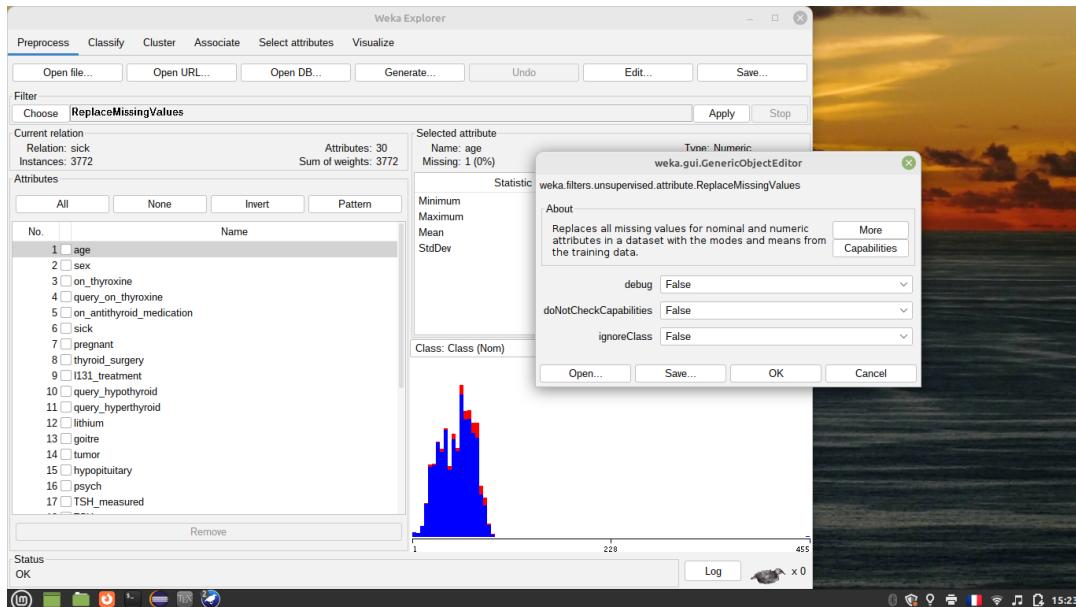


FIGURE 4.7: C4.5 MissingValues

Here we use the **ReplaceMissingValues** filter with our dataset to replace the missing values with the mean value.

### Cross Validation 5 folds

```
==== Run information ====
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues
Instances:   3772
Attributes:  30
             age
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             I131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary
             psych
             TSH_measured
             TSH
             T3_measured
             T3
             TT4_measured
             TT4
```

```

T4U_measured
T4U
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Test mode: 5-fold cross-validation

==== Classifier model (full training set) ===

J48 pruned tree
-----

T3 <= 1.1
| TT4 <= 53
| | TSH <= 19
| | | FTI <= 78
| | | | T4U <= 0.67
| | | | | T3 <= 0.5: sick (3.0/1.0)
| | | | | T3 > 0.5: negative (2.0)
| | | | | T4U > 0.67: sick (7.0)
| | | | | FTI > 78: negative (8.0)
| | | | | TSH > 19: negative (53.0)
| | TT4 > 53
| | | FTI <= 155
| | | | referral_source = SVHC: sick (9.0)
| | | | referral_source = other
| | | | | sick = f
| | | | | age <= 70
| | | | | | TSH <= 9.5: negative (20.0)
| | | | | | TSH > 9.5: sick (6.0/1.0)
| | | | | | age > 70
| | | | | | | on_thyroxine = f: sick (13.0/1.0)
| | | | | | | on_thyroxine = t: negative (3.0/1.0)
| | | | | | | sick = t: sick (3.0)
| | | | | | | referral_source = SVI
| | | | | | | FTI <= 70
| | | | | | | | FTI <= 61: negative (2.0)
| | | | | | | | FTI > 61: sick (8.0)
| | | | | | | | FTI > 70: sick (153.0)
| | | | | | | | referral_source = STMW: sick (0.0)
| | | | | | | | referral_source = SVHD: sick (3.0)
| | | | | | | FTI > 155
| | | | | | | | TT4 <= 92: sick (3.0)
| | | | | | | | TT4 > 92: negative (15.0)

T3 > 1.1
| FTI <= 171
| | TT4 <= 56
| | | referral_source = SVHC: negative (4.0)
| | | referral_source = other: negative (30.0/1.0)
| | | referral_source = SVI
| | | | TSH <= 19: sick (8.0/1.0)
| | | | TSH > 19: negative (5.0)
| | | | referral_source = STMW: negative (2.0)
| | | | referral_source = SVHD: negative (0.0)
| | | TT4 > 56: negative (3281.0/1.0)
| FTI > 171
| | T3 <= 2.6
| | | T3_measured = t
| | | | on_thyroxine = f
| | | | | TSH <= 0.31: sick (18.0/2.0)
| | | | | TSH > 0.31: negative (7.0)
| | | | | on_thyroxine = t: negative (14.0)
| | | | | T3_measured = f: negative (29.0)
| | | | | T3 > 2.6: negative (63.0)

Number of Leaves : 30
Size of the tree : 53

Time taken to build model: 0.14 seconds

==== Stratified cross-validation ===
==== Summary ===

Correctly Classified Instances      3714          98.4624 %
Incorrectly Classified Instances    58           1.5376 %
Kappa statistic                   0.8657
Mean absolute error                0.0178
Root mean squared error            0.1206
Relative absolute error             15.4458 %
Root relative squared error        50.2858 %
Total Number of Instances         3772

```

```
==== Detailed Accuracy By Class ====
      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC     ROC Area  PRC Area  Class
      0.992    0.130    0.992    0.992   0.992    0.866    0.949    0.993    negative
      0.870    0.008    0.878    0.870   0.874    0.866    0.949    0.814    sick
Weighted Avg.  0.985    0.122    0.985    0.985   0.985    0.866    0.949    0.982

==== Confusion Matrix ====
      a     b  <-- classified as
3513  28 |  a = negative
      30  201 |  b = sick
```

## Cross Validation 10 folds

```
==== Run information ====
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   10-fold cross-validation

==== Classifier model (full training set) ====
J48 pruned tree
-----
T3 <= 1.1
|  TT4 <= 53
|  |  TSH <= 19
|  |  |  FTI <= 78
|  |  |  |  T4U <= 0.67
|  |  |  |  |  T3 <= 0.5: sick (3.0/1.0)
|  |  |  |  |  |  T3 > 0.5: negative (2.0)
|  |  |  |  |  |  T4U > 0.67: sick (7.0)
|  |  |  |  |  |  FTI > 78: negative (8.0)
|  |  |  |  |  |  TSH > 19: negative (53.0)
TT4 > 53
|  |  FTI <= 155
|  |  |  referral_source = SVHC: sick (9.0)
|  |  |  |  referral_source = other
|  |  |  |  |  sick = f
|  |  |  |  |  |  age <= 70
|  |  |  |  |  |  |  TSH <= 9.5: negative (20.0)
|  |  |  |  |  |  |  |  TSH > 9.5: sick (6.0/1.0)
|  |  |  |  |  |  |  |  age > 70
|  |  |  |  |  |  |  |  |  on_thyroxine = f: sick (13.0/1.0)
|  |  |  |  |  |  |  |  |  |  on_thyroxine = t: negative (3.0/1.0)
|  |  |  |  |  |  |  |  |  sick = t: sick (3.0)
|  |  |  |  |  |  |  |  |  referral_source = SVI
|  |  |  |  |  |  |  |  |  FTI <= 70
|  |  |  |  |  |  |  |  |  |  FTI <= 61: negative (2.0)
|  |  |  |  |  |  |  |  |  |  FTI > 61: sick (8.0)
```

```

|   |   |   |   FTI > 70: sick (153.0)
|   |   |   referral_source = STMW: sick (0.0)
|   |   |   referral_source = SVHD: sick (3.0)
|   |   FTI > 155
|   |   |   TT4 <= 92: sick (3.0)
|   |   |   TT4 > 92: negative (15.0)
T3 > 1.1
|   FTI <= 171
|   |   TT4 <= 56
|   |   |   referral_source = SVHC: negative (4.0)
|   |   |   referral_source = other: negative (30.0/1.0)
|   |   |   referral_source = SVI
|   |   |   |   TSH <= 19: sick (8.0/1.0)
|   |   |   |   TSH > 19: negative (5.0)
|   |   |   referral_source = STMW: negative (2.0)
|   |   |   referral_source = SVHD: negative (0.0)
|   |   TT4 > 56: negative (3281.0/1.0)
|   FTI > 171
|   |   T3 <= 2.6
|   |   |   T3_measured = t
|   |   |   |   on_thyroxine = f
|   |   |   |   |   TSH <= 0.31: sick (18.0/2.0)
|   |   |   |   |   TSH > 0.31: negative (7.0)
|   |   |   |   on_thyroxine = t: negative (14.0)
|   |   |   T3_measured = f: negative (29.0)
|   |   T3 > 2.6: negative (63.0)

Number of Leaves : 30
Size of the tree : 53

Time taken to build model: 0.13 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3729          98.86 %
Incorrectly Classified Instances    43           1.14 %
Kappa statistic                   0.9002
Mean absolute error               0.0132
Root mean squared error          0.1042
Relative absolute error          11.4806 %
Root relative squared error     43.4437 %
Total Number of Instances         3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.994     0.100     0.994     0.994     0.994     0.900     0.939     0.991     negative
      0.900     0.006     0.912     0.900     0.906     0.900     0.939     0.861     sick
Weighted Avg.      0.989     0.094     0.989     0.989     0.989     0.900     0.939     0.983

==== Confusion Matrix ====

      a     b  <- classified as
3521  20 |    a = negative
      23  208 |    b = sick

```

## Cross Validation 15 folds

```

==== Run information ====

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              i131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured

```

```

TSH
T3_measured
T3
TT4_measured
TT4
T4U_measured
T4U
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Test mode: 15-fold cross-validation

==== Classifier model (full training set) ====
J48 pruned tree
-----
T3 <= 1.1
| TT4 <= 53
| | TSH <= 19
| | | FTI <= 78
| | | | T4U <= 0.67
| | | | | T3 <= 0.5: sick (3.0/1.0)
| | | | | T3 > 0.5: negative (2.0)
| | | | T4U > 0.67: sick (7.0)
| | | | FTI > 78: negative (8.0)
| | | TSH > 19: negative (53.0)
| TT4 > 53
| | FTI <= 155
| | | referral_source = SVHC: sick (9.0)
| | | referral_source = other
| | | | sick = f
| | | | | age <= 70
| | | | | | TSH <= 9.5: negative (20.0)
| | | | | | TSH > 9.5: sick (6.0/1.0)
| | | | | | age > 70
| | | | | | on_thyroxine = f: sick (13.0/1.0)
| | | | | | on_thyroxine = t: negative (3.0/1.0)
| | | | | | sick = t: sick (3.0)
| | | | | referral_source = SVI
| | | | | | FTI <= 70
| | | | | | | FTI <= 61: negative (2.0)
| | | | | | | FTI > 61: sick (8.0)
| | | | | | FTI > 70: sick (153.0)
| | | | | referral_source = STMW: sick (0.0)
| | | | | referral_source = SVHD: sick (3.0)
| | | | FTI > 155
| | | | | TT4 <= 92: sick (3.0)
| | | | | TT4 > 92: negative (15.0)
T3 > 1.1
| FTI <= 171
| | TT4 <= 56
| | | referral_source = SVHC: negative (4.0)
| | | referral_source = other: negative (30.0/1.0)
| | | referral_source = SVI
| | | | TSH <= 19: sick (8.0/1.0)
| | | | | TSH > 19: negative (5.0)
| | | | referral_source = STMW: negative (2.0)
| | | | referral_source = SVHD: negative (0.0)
| | | TT4 > 56: negative (3281.0/1.0)
| FTI > 171
| | T3 <= 2.6
| | | T3_measured = t
| | | | on_thyroxine = f
| | | | | TSH <= 0.31: sick (18.0/2.0)
| | | | | TSH > 0.31: negative (7.0)
| | | | | on_thyroxine = t: negative (14.0)
| | | | T3_measured = f: negative (29.0)
| | | | T3 > 2.6: negative (63.0)

Number of Leaves : 30
Size of the tree : 53

Time taken to build model: 0.12 seconds

==== Stratified cross-validation ====
==== Summary ====

```

	Correctly Classified Instances	98.754 %
Incorrectly Classified Instances	47	1.246 %
Kappa statistic	0.8896	
Mean absolute error	0.015	

```

Root mean squared error          0.1078
Relative absolute error        13.037 %
Root relative squared error    44.9396 %
Total Number of Instances      3772

==== Detailed Accuracy By Class ====

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC     ROC Area  PRC Area  Class
      0.995    0.121    0.992    0.995    0.993    0.890    0.955    0.992    negative
      0.879    0.005    0.914    0.879    0.896    0.890    0.955    0.879    sick
Weighted Avg.      0.988    0.114    0.987    0.988    0.987    0.890    0.955    0.985

==== Confusion Matrix ====
a      b  <-- classified as
3522  19 |  a = negative
 28  203 |  b = sick

```

## Percentage Split 50%

```

==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              t131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 50.0% train, remainder test

==== Classifier model (full training set) ===

J48 pruned tree
-----
T3 <= 1.1
|  TT4 <= 53
|  |  TSH <= 19
|  |  |  FTI <= 78
|  |  |  |  T4U <= 0.67
|  |  |  |  |  T3 <= 0.5: sick (3.0/1.0)
|  |  |  |  |  |  T3 > 0.5: negative (2.0)
|  |  |  |  |  |  T4U > 0.67: sick (7.0)
|  |  |  |  FTI > 78: negative (8.0)
|  |  TSH > 19: negative (53.0)
|  TT4 > 53
|  |  FTI <= 155
|  |  |  referral_source = SVHC: sick (9.0)
|  |  |  referral_source = other
|  |  |  |  sick = f
|  |  |  |  |  age <= 70
|  |  |  |  |  |  TSH <= 9.5: negative (20.0)
|  |  |  |  |  |  TSH > 9.5: sick (6.0/1.0)
|  |  |  |  |  |  age > 70
|  |  |  |  |  |  |  on_thyroxine = f: sick (13.0/1.0)
|  |  |  |  |  |  |  on_thyroxine = t: negative (3.0/1.0)

```

```

|   |   |   |   sick = t: sick (3.0)
|   |   |   referral_source = SVI
|   |   |   |   FTI <= 70
|   |   |   |   |   FTI <= 61: negative (2.0)
|   |   |   |   |   FTI > 61: sick (8.0)
|   |   |   |   FTI > 70: sick (153.0)
|   |   |   referral_source = STMW: sick (0.0)
|   |   |   referral_source = SVHD: sick (3.0)
|   |   FTI > 155
|   |   |   TT4 <= 92: sick (3.0)
|   |   |   TT4 > 92: negative (15.0)
T3 > 1.1
|   FTI <= 171
|   |   TT4 <= 56
|   |   |   referral_source = SVHC: negative (4.0)
|   |   |   referral_source = other: negative (30.0/1.0)
|   |   |   referral_source = SVI
|   |   |   |   TSH <= 19: sick (8.0/1.0)
|   |   |   |   TSH > 19: negative (5.0)
|   |   |   referral_source = STMW: negative (2.0)
|   |   |   referral_source = SVHD: negative (0.0)
|   |   TT4 > 56: negative (3281.0/1.0)
|   FTI > 171
|   |   T3 <= 2.6
|   |   |   T3_measured = t
|   |   |   |   on_thyroxine = f
|   |   |   |   |   TSH <= 0.31: sick (18.0/2.0)
|   |   |   |   |   TSH > 0.31: negative (7.0)
|   |   |   |   on_thyroxine = t: negative (14.0)
|   |   |   T3_measured = f: negative (29.0)
|   |   T3 > 2.6: negative (63.0)

Number of Leaves : 30
Size of the tree : 53

Time taken to build model: 0.12 seconds
== Evaluation on test split ==
Time taken to test model on test split: 0.01 seconds
== Summary ==
Correctly Classified Instances      1853          98.2503 %
Incorrectly Classified Instances    33           1.7497 %
Kappa statistic                      0.8387
Mean absolute error                  0.0246
Root mean squared error              0.1262
Relative absolute error              21.3425 %
Root relative squared error         52.3352 %
Total Number of Instances            1886

== Detailed Accuracy By Class ==

      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC     ROC Area  PRC Area  Class
0.995     0.214     0.986     0.995     0.991     0.842     0.916     0.989  negative
0.786     0.005     0.920     0.786     0.848     0.842     0.916     0.797  sick
Weighted Avg.      0.983     0.201     0.982     0.983     0.982     0.842     0.916     0.977

== Confusion Matrix ==

      a      b  <- classified as
1761    8 |    a = negative
      25   92 |    b = sick

```

## Percentage Split 60%

```

== Run information ==
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              T131_treatment

```

```

query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH
T3_measured
T3
TT4_measured
TT4
T4U_measured
T4U
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Test mode: split 60.0% train, remainder test

==== Classifier model (full training set) ====

J48 pruned tree
-----
T3 <= 1.1
|   TT4 <= 53
|   |   TSH <= 19
|   |   |   FTI <= 78
|   |   |   |   T4U <= 0.67
|   |   |   |   |   T3 <= 0.5: sick (3.0/1.0)
|   |   |   |   |   |   T3 > 0.5: negative (2.0)
|   |   |   |   |   T4U > 0.67: sick (7.0)
|   |   |   |   FTI > 78: negative (8.0)
|   |   |   TSH > 19: negative (53.0)
|   TT4 > 53
|   |   FTI <= 155
|   |   |   referral_source = SVHC: sick (9.0)
|   |   |   referral_source = other
|   |   |   |   sick = f
|   |   |   |   age <= 70
|   |   |   |   |   TSH <= 9.5: negative (20.0)
|   |   |   |   |   |   TSH > 9.5: sick (6.0/1.0)
|   |   |   |   |   age > 70
|   |   |   |   |   |   on_thyroxine = f: sick (13.0/1.0)
|   |   |   |   |   |   on_thyroxine = t: negative (3.0/1.0)
|   |   |   |   |   sick = t: sick (3.0)
|   |   |   |   referral_source = SVI
|   |   |   |   |   FTI <= 70
|   |   |   |   |   |   FTI <= 61: negative (2.0)
|   |   |   |   |   |   |   FTI > 61: sick (8.0)
|   |   |   |   |   |   FTI > 70: sick (153.0)
|   |   |   |   |   referral_source = STMW: sick (0.0)
|   |   |   |   |   referral_source = SVHD: sick (3.0)
|   |   |   |   FTI > 155
|   |   |   |   |   TT4 <= 92: sick (3.0)
|   |   |   |   |   TT4 > 92: negative (15.0)
T3 > 1.1
|   FTI <= 171
|   |   TT4 <= 56
|   |   |   referral_source = SVHC: negative (4.0)
|   |   |   referral_source = other: negative (30.0/1.0)
|   |   |   referral_source = SVI
|   |   |   |   TSH <= 19: sick (8.0/1.0)
|   |   |   |   |   TSH > 19: negative (5.0)
|   |   |   |   referral_source = STMW: negative (2.0)
|   |   |   |   referral_source = SVHD: negative (0.0)
|   |   |   TT4 > 56: negative (3281.0/1.0)
|   FTI > 171
|   |   T3 <= 2.6
|   |   |   T3_measured = t
|   |   |   |   on_thyroxine = f
|   |   |   |   |   TSH <= 0.31: sick (18.0/2.0)
|   |   |   |   |   |   TSH > 0.31: negative (7.0)
|   |   |   |   |   on_thyroxine = t: negative (14.0)
|   |   |   |   T3_measured = f: negative (29.0)
|   |   |   T3 > 2.6: negative (63.0)

Number of Leaves : 30
Size of the tree : 53

Time taken to build model: 0.12 seconds

```

```

==== Evaluation on test split ====
Time taken to test model on test split: 0.01 seconds
==== Summary ====
Correctly Classified Instances      1486          98.4758 %
Incorrectly Classified Instances    23           1.5242 %
Kappa statistic                   0.8474
Mean absolute error               0.0227
Root mean squared error          0.1201
Relative absolute error          19.7702 %
Root relative squared error     50.9757 %
Total Number of Instances        1509

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.999     0.236     0.985      0.999    0.992      0.854     0.887     0.987     negative
      0.764     0.001     0.971      0.764    0.855      0.854     0.887     0.810     sick
Weighted Avg.       0.985     0.222     0.985      0.985    0.984      0.854     0.887     0.976

==== Confusion Matrix ====

      a      b  <-- classified as
1418    2 |    a = negative
      21   68 |    b = sick

```

## Percentage Split 66%

```

==== Run information ====
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues
Instances:   3772
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 66.0% train, remainder test

==== Classifier model (full training set) ====
J48 pruned tree
-----
T3 <= 1.1
|   TT4 <= 53
|   |   TSH <= 19
|   |   |   FTI <= 78
|   |   |   |   T4U <= 0.67
|   |   |   |   |   T3 <= 0.5: sick (3.0/1.0)
|   |   |   |   |   |   T3 > 0.5: negative (2.0)
|   |   |   |   |   T4U > 0.67: sick (7.0)
|   |   |   |   FTI > 78: negative (8.0)
|   |   TSH > 19: negative (53.0)

```

```

|   TT4 > 53
|   |   FTI <= 155
|   |   |   referral_source = SVHC: sick (9.0)
|   |   |   referral_source = other
|   |   |   |   sick = f
|   |   |   |   age <= 70
|   |   |   |   |   TSH <= 9.5: negative (20.0)
|   |   |   |   |   TSH > 9.5: sick (6.0/1.0)
|   |   |   |   age > 70
|   |   |   |   |   on_thyroxine = f: sick (13.0/1.0)
|   |   |   |   |   on_thyroxine = t: negative (3.0/1.0)
|   |   |   |   sick = t: sick (3.0)
|   |   |   referral_source = SVI
|   |   |   |   FTI <= 70
|   |   |   |   |   FTI <= 61: negative (2.0)
|   |   |   |   |   FTI > 61: sick (8.0)
|   |   |   |   FTI > 70: sick (153.0)
|   |   |   referral_source = STMW: sick (0.0)
|   |   |   referral_source = SVHD: sick (3.0)
|   |   FTI > 155
|   |   |   TT4 <= 92: sick (3.0)
|   |   |   TT4 > 92: negative (15.0)
T3 > 1.1
|   FTI <= 171
|   |   TT4 <= 56
|   |   |   referral_source = SVHC: negative (4.0)
|   |   |   referral_source = other: negative (30.0/1.0)
|   |   |   referral_source = SVI
|   |   |   |   TSH <= 19: sick (8.0/1.0)
|   |   |   |   TSH > 19: negative (5.0)
|   |   |   referral_source = STMW: negative (2.0)
|   |   |   referral_source = SVHD: negative (0.0)
|   |   TT4 > 56: negative (3281.0/1.0)
|   FTI > 171
|   |   T3 <= 2.6
|   |   |   T3_measured = t
|   |   |   |   on_thyroxine = f
|   |   |   |   |   TSH <= 0.31: sick (18.0/2.0)
|   |   |   |   |   TSH > 0.31: negative (7.0)
|   |   |   |   on_thyroxine = t: negative (14.0)
|   |   |   T3_measured = f: negative (29.0)
|   |   T3 > 2.6: negative (63.0)

Number of Leaves : 30
Size of the tree : 53

Time taken to build model: 0.1 seconds
== Evaluation on test split ==
Time taken to test model on test split: 0 seconds
== Summary ==
Correctly Classified Instances      1266          98.752 %
Incorrectly Classified Instances     16           1.248 %
Kappa statistic                      0.8791
Mean absolute error                  0.0154
Root mean squared error              0.1095
Relative absolute error              13.4536 %
Root relative squared error         46.6573 %
Total Number of Instances           1282

== Detailed Accuracy By Class ==


|               | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC   | ROC Area | PRC Area | Class |
|---------------|---------|---------|-----------|--------|-----------|-------|----------|----------|-------|
| 0.998         | 0.173   | 0.989   | 0.998     | 0.993  | 0.882     | 0.906 | 0.989    | negative |       |
| 0.827         | 0.002   | 0.954   | 0.827     | 0.886  | 0.882     | 0.906 | 0.849    | sick     |       |
| Weighted Avg. | 0.988   | 0.163   | 0.987     | 0.988  | 0.987     | 0.882 | 0.906    | 0.981    |       |


== Confusion Matrix ==
    a      b  <-- classified as
1204    3 |    a = negative
    13   62 |    b = sick

```

## Leave one out fold

```

== Run information ==
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues

```

```

Instances:      3772
Attributes:    30
               age
               sex
               on_thyroxine
               query_on_thyroxine
               on_antithyroid_medication
               sick
               pregnant
               thyroid_surgery
               I131_treatment
               query_hypothyroid
               query_hyperthyroid
               lithium
               goitre
               tumor
               hypopituitary
               psych
               TSH_measured
               TSH
               T3_measured
               T3
               TT4_measured
               TT4
               T4U_measured
               T4U
               FTI_measured
               FTI
               TBG_measured
               TBG
               referral_source
               Class
Test mode:    3772-fold cross-validation

==== Classifier model (full training set) ====

J48 pruned tree
-----
T3 <= 1.1
|   TT4 <= 53
|   |   TSH <= 19
|   |   |   FTI <= 78
|   |   |   |   T4U <= 0.67
|   |   |   |   |   T3 <= 0.5: sick (3.0/1.0)
|   |   |   |   |   |   T3 > 0.5: negative (2.0)
|   |   |   |   |   T4U > 0.67: sick (7.0)
|   |   |   |   FTI > 78: negative (8.0)
|   |   TSH > 19: negative (53.0)
|   TT4 > 53
|   |   FTI <= 155
|   |   |   referral_source = SVHC: sick (9.0)
|   |   |   referral_source = other
|   |   |   |   sick = f
|   |   |   |   age <= 70
|   |   |   |   |   TSH <= 9.5: negative (20.0)
|   |   |   |   |   TSH > 9.5: sick (6.0/1.0)
|   |   |   |   |   age > 70
|   |   |   |   |   on_thyroxine = f: sick (13.0/1.0)
|   |   |   |   |   on_thyroxine = t: negative (3.0/1.0)
|   |   |   |   |   sick = t: sick (3.0)
|   |   |   |   referral_source = SVI
|   |   |   |   FTI <= 70
|   |   |   |   |   FTI <= 61: negative (2.0)
|   |   |   |   |   FTI > 61: sick (8.0)
|   |   |   |   |   FTI > 70: sick (153.0)
|   |   |   |   referral_source = STMW: sick (0.0)
|   |   |   |   referral_source = SVHD: sick (3.0)
|   |   |   FTI > 155
|   |   |   |   TT4 <= 92: sick (3.0)
|   |   |   |   TT4 > 92: negative (15.0)
T3 > 1.1
|   FTI <= 171
|   |   TT4 <= 56
|   |   |   referral_source = SVHC: negative (4.0)
|   |   |   referral_source = other: negative (30.0/1.0)
|   |   |   referral_source = SVI
|   |   |   |   TSH <= 19: sick (8.0/1.0)
|   |   |   |   TSH > 19: negative (5.0)
|   |   |   |   referral_source = STMW: negative (2.0)
|   |   |   |   referral_source = SVHD: negative (0.0)
|   |   |   TT4 > 56: negative (3281.0/1.0)
|   |   FTI > 171
|   |   |   T3 <= 2.6
|   |   |   |   T3_measured = t
|   |   |   |   |   on_thyroxine = f
|   |   |   |   |   |   TSH <= 0.31: sick (18.0/2.0)

```

```

|   |   |   |   TSH > 0.31: negative (7.0)
|   |   |   on_thyroxine = t: negative (14.0)
|   |   |   T3_measured = f: negative (29.0)
|   |   T3 > 2.6: negative (63.0)

Number of Leaves : 30

Size of the tree : 53

Time taken to build model: 0.12 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3722          98.6744 %
Incorrectly Classified Instances    50           1.3256 %
Kappa statistic                   0.8809
Mean absolute error               0.0137
Root mean squared error          0.1099
Relative absolute error          11.8506 %
Root relative squared error     45.813 %
Total Number of Instances        3772

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.995     0.143     0.991     0.995     0.993     0.881     0.948     0.994     negative
      0.857     0.005     0.921     0.857     0.888     0.881     0.948     0.903     sick
Weighted Avg.      0.987     0.134     0.986     0.987     0.987     0.881     0.948     0.988

==== Confusion Matrix ====

      a     b  <-- classified as
3524  17 |     a = negative
      33 198 |     b = sick

```

## 4.5.2 Extreme Values

### Cross Validation 5 folds

```

==== Run information ====

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3627
Attributes:  30
             age
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             I131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary
             psych
             TSH_measured
             TSH
             T3_measured
             T3
             TT4_measured
             TT4
             T4U_measured
             T4U
             FTI_measured
             FTI
             TBG_measured
             TBG
             referral_source
             Class
Test mode:   5-fold cross-validation

==== Classifier model (full training set) ====

J48 pruned tree
-----

```

```

T3 <= 1.1
| FTI <= 155
| | referral_source = SVHC: sick (9.0)
| | referral_source = other
| | | sick = f
| | | | age <= 70: negative (23.0/1.0)
| | | | age > 70
| | | | | T4U <= 0.95: sick (11.0)
| | | | | T4U > 0.95: negative (4.0/1.0)
| | | | sick = t: sick (3.0)
| | | referral_source = SVI
| | | | TT4 <= 53
| | | | | sex = F
| | | | | | age <= 55: sick (3.0)
| | | | | | age > 55: negative (10.0/1.0)
| | | | | sex = M
| | | | | | age <= 51: negative (4.0/1.0)
| | | | | | age > 51: sick (4.0)
| | | | TT4 > 53
| | | | | FTI <= 70
| | | | | | FTI <= 61: negative (2.0)
| | | | | | FTI > 61: sick (8.0)
| | | | | FTI > 70: sick (153.0)
| | | | referral_source = STMW: sick (0.0)
| | | | referral_source = SVHD: sick (3.0)
| | FTI > 155
| | | TT4 <= 92
| | | | TSH <= 0.5: negative (2.0)
| | | | TSH > 0.5: sick (3.0)
| | | TT4 > 92: negative (15.0)

T3 > 1.1
| FTI <= 171
| | TT4 <= 56
| | | referral_source = SVHC: negative (4.0)
| | | referral_source = other: negative (15.0/1.0)
| | | referral_source = SVI: sick (8.0/1.0)
| | | referral_source = STMW: negative (1.0)
| | | referral_source = SVHD: negative (0.0)
| | | TT4 > 56: negative (3230.0/1.0)
| | FTI > 171
| | | on_thyroxine = f
| | | | T3 <= 2.7
| | | | | T3_measured = t
| | | | | | TSH <= 0.31: sick (18.0/2.0)
| | | | | | TSH > 0.31: negative (7.0)
| | | | | | T3_measured = f: negative (15.0)
| | | | | T3 > 2.7: negative (38.0)
| | | | on_thyroxine = t: negative (34.0)

Number of Leaves : 28
Size of the tree : 49

Time taken to build model: 0.43 seconds

==== Stratified cross-validation ====
==== Summary ====

    Correctly Classified Instances      3575          98.5663 %
    Incorrectly Classified Instances     52           1.4337 %
    Kappa statistic                      0.8768
    Mean absolute error                  0.0159
    Root mean squared error              0.1152
    Relative absolute error             13.5778 %
    Root relative squared error         47.6675 %
    Total Number of Instances            3627

==== Detailed Accuracy By Class ====

          TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
          0.993     0.119     0.992     0.993     0.992     0.877     0.966     0.995   negative
          0.881     0.007     0.888     0.881     0.884     0.877     0.966     0.821   sick
Weighted Avg.   0.986     0.112     0.986     0.986     0.986     0.877     0.966     0.985

==== Confusion Matrix ====

      a     b  <- classified as
3376  25 | a = negative
      27  199 | b = sick

```

## Cross Validation 10 folds

```

==== Run information ====

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3627
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   10-fold cross-validation

==== Classifier model (full training set) ====

J48 pruned tree
-----
T3 <= 1.1
|  FTI <= 155
|  |  referral_source = SVHC: sick (9.0)
|  |  referral_source = other
|  |  |  sick = f
|  |  |  |  age <= 70: negative (23.0/1.0)
|  |  |  |  age > 70
|  |  |  |  |  T4U <= 0.95: sick (11.0)
|  |  |  |  |  |  T4U > 0.95: negative (4.0/1.0)
|  |  |  |  sick = t: sick (3.0)
|  |  |  referral_source = SVI
|  |  |  TT4 <= 53
|  |  |  |  sex = F
|  |  |  |  |  age <= 55: sick (3.0)
|  |  |  |  |  age > 55: negative (10.0/1.0)
|  |  |  |  sex = M
|  |  |  |  |  age <= 51: negative (4.0/1.0)
|  |  |  |  |  age > 51: sick (4.0)
|  |  |  |  TT4 > 53
|  |  |  |  |  FTI <= 70
|  |  |  |  |  |  FTI <= 61: negative (2.0)
|  |  |  |  |  |  FTI > 61: sick (8.0)
|  |  |  |  |  FTI > 70: sick (153.0)
|  |  |  |  referral_source = STMW: sick (0.0)
|  |  |  |  referral_source = SVHD: sick (3.0)
|  |  |  FTI > 155
|  |  |  |  TT4 <= 92
|  |  |  |  |  TSH <= 0.5: negative (2.0)
|  |  |  |  |  TSH > 0.5: sick (3.0)
|  |  |  |  TT4 > 92: negative (15.0)
T3 > 1.1
|  FTI <= 171
|  |  TT4 <= 56
|  |  |  referral_source = SVHC: negative (4.0)
|  |  |  referral_source = other: negative (15.0/1.0)
|  |  |  referral_source = SVI: sick (8.0/1.0)
|  |  |  referral_source = STMW: negative (1.0)
|  |  |  referral_source = SVHD: negative (0.0)
|  |  |  TT4 > 56: negative (3230.0/1.0)
|  |  FTI > 171
|  |  |  on_thyroxine = f
|  |  |  |  T3 <= 2.7

```

```

|   |   |   |   T3_measured = t
|   |   |   |   |   TSH <= 0.31: sick (18.0/2.0)
|   |   |   |   |   |   TSH > 0.31: negative (7.0)
|   |   |   |   T3_measured = f: negative (15.0)
|   |   |   T3 > 2.7: negative (38.0)
|   |   on_thyroxine = t: negative (34.0)

Number of Leaves : 28

Size of the tree : 49

Time taken to build model: 0.13 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3584          98.8144 %
Incorrectly Classified Instances    43           1.1856 %
Kappa statistic                      0.8953
Mean absolute error                  0.0142
Root mean squared error              0.1068
Relative absolute error              12.0943 %
Root relative squared error         44.1779 %
Total Number of Instances           3627

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC     ROC Area   PRC Area   Class
      0.996     0.128     0.992     0.996     0.994     0.896     0.941     0.991   negative
      0.872     0.004     0.934     0.872     0.902     0.896     0.941     0.875   sick
Weighted Avg.      0.988     0.121     0.988     0.988     0.988     0.896     0.941     0.984

==== Confusion Matrix ====

      a     b  <-- classified as
3387   14 |     a = negative
      29  197 |     b = sick

```

## Cross Validation 15 folds

```

==== Run information ====

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3627
Attributes:  30
             age
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             I131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary
             psych
             TSH_measured
             TSH
             T3_measured
             T3
             TT4_measured
             TT4
             T4U_measured
             T4U
             FTI_measured
             FTI
             TBG_measured
             TBG
             referral_source
             Class
Test mode:   15-fold cross-validation

==== Classifier model (full training set) ====
J48 pruned tree
-----
```

```

T3 <= 1.1
|   FTI <= 155
|   |   referral_source = SVHC: sick (9.0)
|   |   referral_source = other
|   |   |   sick = f
|   |   |   age <= 70: negative (23.0/1.0)
|   |   |   age > 70
|   |   |   |   TAU <= 0.95: sick (11.0)
|   |   |   |   TAU > 0.95: negative (4.0/1.0)
|   |   |   sick = t: sick (3.0)
|   |   referral_source = SVI
|   |   TT4 <= 53
|   |   |   sex = F
|   |   |   |   age <= 55: sick (3.0)
|   |   |   |   age > 55: negative (10.0/1.0)
|   |   |   sex = M
|   |   |   |   age <= 51: negative (4.0/1.0)
|   |   |   |   age > 51: sick (4.0)
|   |   TT4 > 53
|   |   |   FTI <= 70
|   |   |   |   FTI <= 61: negative (2.0)
|   |   |   |   FTI > 61: sick (8.0)
|   |   |   FTI > 70: sick (153.0)
|   |   referral_source = STMW: sick (0.0)
|   |   referral_source = SVHD: sick (3.0)
|   FTI > 155
|   TT4 <= 92
|   |   TSH <= 0.5: negative (2.0)
|   |   TSH > 0.5: sick (3.0)
|   TT4 > 92: negative (15.0)

T3 > 1.1
|   FTI <= 171
|   TT4 <= 56
|   |   referral_source = SVHC: negative (4.0)
|   |   referral_source = other: negative (15.0/1.0)
|   |   referral_source = SVI: sick (8.0/1.0)
|   |   referral_source = STMW: negative (1.0)
|   |   referral_source = SVHD: negative (0.0)
|   |   TT4 > 56: negative (3230.0/1.0)
|   FTI > 171
|   |   on_thyroxine = f
|   |   |   T3 <= 2.7
|   |   |   |   T3_measured = t
|   |   |   |   |   TSH <= 0.31: sick (18.0/2.0)
|   |   |   |   |   TSH > 0.31: negative (7.0)
|   |   |   |   T3_measured = f: negative (15.0)
|   |   |   T3 > 2.7: negative (38.0)
|   |   on_thyroxine = t: negative (34.0)

Number of Leaves : 28
Size of the tree : 49

Time taken to build model: 0.1 seconds

==== Stratified cross-validation ====
==== Summary ====

    Correctly Classified Instances      3594          99.0902 %
    Incorrectly Classified Instances     33           0.9098 %
    Kappa statistic                      0.9217
    Mean absolute error                  0.0103
    Root mean squared error              0.0922
    Relative absolute error              8.8101 %
    Root relative squared error         38.1374 %
    Total Number of Instances            3627

==== Detailed Accuracy By Class ====

          TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
          0.996     0.080     0.995     0.996     0.995     0.922     0.956     0.994     negative
          0.920     0.004     0.933     0.920     0.927     0.922     0.956     0.903     sick
Weighted Avg.    0.991     0.075     0.991     0.991     0.991     0.922     0.956     0.988

==== Confusion Matrix ====

      a     b  <- classified as
3386  15 |     a = negative
      18  208 |     b = sick

```

### Percentage Split 50%

```
==== Run information ====
```

```

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3627
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              t131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 50.0% train, remainder test

==== Classifier model (full training set) ===

J48 pruned tree
-----

T3 <= 1.1
|  FTI <= 155
|  |  referral_source = SVHC: sick (9.0)
|  |  referral_source = other
|  |  |  sick = f
|  |  |  |  age <= 70: negative (23.0/1.0)
|  |  |  |  age > 70
|  |  |  |  |  T4U <= 0.95: sick (11.0)
|  |  |  |  |  T4U > 0.95: negative (4.0/1.0)
|  |  |  |  sick = t: sick (3.0)
|  |  |  |  referral_source = SVI
|  |  |  |  TT4 <= 53
|  |  |  |  |  sex = F
|  |  |  |  |  |  age <= 55: sick (3.0)
|  |  |  |  |  |  age > 55: negative (10.0/1.0)
|  |  |  |  |  sex = M
|  |  |  |  |  |  age <= 51: negative (4.0/1.0)
|  |  |  |  |  |  age > 51: sick (4.0)
|  |  |  |  TT4 > 53
|  |  |  |  |  FTI <= 70
|  |  |  |  |  |  FTI <= 61: negative (2.0)
|  |  |  |  |  |  FTI > 61: sick (8.0)
|  |  |  |  |  FTI > 70: sick (153.0)
|  |  |  |  |  referral_source = STMW: sick (0.0)
|  |  |  |  |  referral_source = SVHD: sick (3.0)
|  FTI > 155
|  |  TT4 <= 92
|  |  |  TSH <= 0.5: negative (2.0)
|  |  |  TSH > 0.5: sick (3.0)
|  |  TT4 > 92: negative (15.0)

T3 > 1.1
|  FTI <= 171
|  |  TT4 <= 56
|  |  |  referral_source = SVHC: negative (4.0)
|  |  |  referral_source = other: negative (15.0/1.0)
|  |  |  referral_source = SVI: sick (8.0/1.0)
|  |  |  referral_source = STMW: negative (1.0)
|  |  |  referral_source = SVHD: negative (0.0)
|  |  TT4 > 56: negative (3230.0/1.0)
|  FTI > 171
|  |  on_thyroxine = f
|  |  |  T3 <= 2.7
|  |  |  |  T3_measured = t
|  |  |  |  |  TSH <= 0.31: sick (18.0/2.0)
|  |  |  |  |  TSH > 0.31: negative (7.0)

```

```

|   |   |   |   T3_measured = f: negative (15.0)
|   |   |   T3 > 2.7: negative (38.0)
|   |   on_thyroxine = t: negative (34.0)

Number of Leaves : 28

Size of the tree : 49

Time taken to build model: 0.12 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0.02 seconds

==== Summary ===

Correctly Classified Instances      1783          98.3453 %
Incorrectly Classified Instances    30           1.6547 %
Kappa statistic                      0.8484
Mean absolute error                  0.0179
Root mean squared error              0.1198
Relative absolute error              15.2589 %
Root relative squared error         51.0446 %
Total Number of Instances           1813

==== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC     ROC Area  PRC Area  Class
      0.992    0.151    0.991    0.992    0.991    0.848    0.883    0.986    negative
      0.849    0.008    0.865    0.849    0.857    0.848    0.883    0.810    sick
Weighted Avg.  0.983    0.143    0.983    0.983    0.983    0.848    0.883    0.976

==== Confusion Matrix ===

      a      b  <-- classified as
1693   14 |      a = negative
      16   90 |      b = sick

```

## Percentage Split 60%

```

==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3627
Attributes:  30
             age
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             I131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary
             psych
             TSH_measured
             TSH
             T3_measured
             T3
             TT4_measured
             TT4
             T4U_measured
             T4U
             FTI_measured
             FTI
             TBG_measured
             TBG
             referral_source
             Class
Test mode:   split 60.0% train, remainder test

==== Classifier model (full training set) ===

J48 pruned tree
-----

```

```

T3 <= 1.1
|   FTI <= 155
|   |   referral_source = SVHC: sick (9.0)
|   |   referral_source = other
|   |   |   sick = f
|   |   |   age <= 70: negative (23.0/1.0)
|   |   |   age > 70
|   |   |   |   T4U <= 0.95: sick (11.0)
|   |   |   |   T4U > 0.95: negative (4.0/1.0)
|   |   |   sick = t: sick (3.0)
|   |   referral_source = SVI
|   |   TT4 <= 53
|   |   |   sex = F
|   |   |   |   age <= 55: sick (3.0)
|   |   |   |   age > 55: negative (10.0/1.0)
|   |   |   sex = M
|   |   |   |   age <= 51: negative (4.0/1.0)
|   |   |   |   age > 51: sick (4.0)
|   |   TT4 > 53
|   |   |   FTI <= 70
|   |   |   |   FTI <= 61: negative (2.0)
|   |   |   |   FTI > 61: sick (8.0)
|   |   |   FTI > 70: sick (153.0)
|   |   referral_source = STMW: sick (0.0)
|   |   referral_source = SVHD: sick (3.0)
|   FTI > 155
|   |   TT4 <= 92
|   |   |   TSH <= 0.5: negative (2.0)
|   |   |   TSH > 0.5: sick (3.0)
|   |   TT4 > 92: negative (15.0)

T3 > 1.1
|   FTI <= 171
|   |   TT4 <= 56
|   |   |   referral_source = SVHC: negative (4.0)
|   |   |   referral_source = other: negative (15.0/1.0)
|   |   |   referral_source = SVI: sick (8.0/1.0)
|   |   |   referral_source = STMW: negative (1.0)
|   |   |   referral_source = SVHD: negative (0.0)
|   |   TT4 > 56: negative (3230.0/1.0)
|   FTI > 171
|   |   on_thyroxine = f
|   |   |   T3 <= 2.7
|   |   |   |   T3_measured = t
|   |   |   |   |   TSH <= 0.31: sick (18.0/2.0)
|   |   |   |   |   TSH > 0.31: negative (7.0)
|   |   |   |   T3_measured = f: negative (15.0)
|   |   |   T3 > 2.7: negative (38.0)
|   |   on_thyroxine = t: negative (34.0)

Number of Leaves : 28
Size of the tree : 49

Time taken to build model: 0.11 seconds
== Evaluation on test split ==
Time taken to test model on test split: 0.03 seconds
== Summary ==
Correctly Classified Instances      1436          98.9662 %
Incorrectly Classified Instances    15           1.0338 %
Kappa statistic                      0.899
Mean absolute error                  0.0126
Root mean squared error              0.0993
Relative absolute error              10.8847 %
Root relative squared error         42.9434 %
Total Number of Instances           1451

== Detailed Accuracy By Class ==
             TP Rate  FP Rate  Precision  Recall   F-Measure  MCC     ROC Area  PRC Area  Class
              0.997    0.134    0.992    0.997    0.995    0.900    0.944    0.993  negative
              0.866    0.003    0.947    0.866    0.904    0.900    0.944    0.855   sick
Weighted Avg.    0.990    0.127    0.989    0.990    0.989    0.900    0.944    0.986

== Confusion Matrix ==
      a      b  <- classified as
1365    4 |  a = negative
      11   71 |  b = sick

```

## Percentage Split 66%

```

==== Run information ====

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3627
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 66.0% train, remainder test

==== Classifier model (full training set) ====

J48 pruned tree
-----
T3 <= 1.1
|  FTI <= 155
|  |  referral_source = SVHC: sick (9.0)
|  |  referral_source = other
|  |  |  sick = f
|  |  |  |  age <= 70: negative (23.0/1.0)
|  |  |  |  age > 70
|  |  |  |  |  T4U <= 0.95: sick (11.0)
|  |  |  |  |  T4U > 0.95: negative (4.0/1.0)
|  |  |  |  sick = t: sick (3.0)
|  |  |  referral_source = SVI
|  |  |  TT4 <= 53
|  |  |  |  sex = F
|  |  |  |  |  age <= 55: sick (3.0)
|  |  |  |  |  age > 55: negative (10.0/1.0)
|  |  |  |  sex = M
|  |  |  |  |  age <= 51: negative (4.0/1.0)
|  |  |  |  |  age > 51: sick (4.0)
|  |  |  |  TT4 > 53
|  |  |  |  |  FTI <= 70
|  |  |  |  |  |  FTI <= 61: negative (2.0)
|  |  |  |  |  |  FTI > 61: sick (8.0)
|  |  |  |  |  FTI > 70: sick (153.0)
|  |  |  |  referral_source = STMW: sick (0.0)
|  |  |  |  referral_source = SVHD: sick (3.0)
|  |  |  FTI > 155
|  |  |  |  TT4 <= 92
|  |  |  |  |  TSH <= 0.5: negative (2.0)
|  |  |  |  |  TSH > 0.5: sick (3.0)
|  |  |  |  TT4 > 92: negative (15.0)
T3 > 1.1
|  FTI <= 171
|  |  TT4 <= 56
|  |  |  referral_source = SVHC: negative (4.0)
|  |  |  referral_source = other: negative (15.0/1.0)
|  |  |  referral_source = SVI: sick (8.0/1.0)
|  |  |  referral_source = STMW: negative (1.0)
|  |  |  referral_source = SVHD: negative (0.0)
|  |  |  TT4 > 56: negative (3230.0/1.0)
|  |  FTI > 171
|  |  |  on_thyroxine = f
|  |  |  |  T3 <= 2.7

```

```

|   |   |   |   T3_measured = t
|   |   |   |   TSH <= 0.31: sick (18.0/2.0)
|   |   |   |   TSH > 0.31: negative (7.0)
|   |   |   T3_measured = f: negative (15.0)
|   |   T3 > 2.7: negative (38.0)
|   |   on_thyroxine = t: negative (34.0)

Number of Leaves : 28

Size of the tree : 49

Time taken to build model: 0.12 seconds

==== Evaluation on test split ====

Time taken to test model on test split: 0.01 seconds

==== Summary ====

Correctly Classified Instances      1220          98.9457 %
Incorrectly Classified Instances    13           1.0543 %
Kappa statistic                   0.9035
Mean absolute error               0.0113
Root mean squared error          0.0981
Relative absolute error          9.7091 %
Root relative squared error     41.0443 %
Total Number of Instances        1233

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
      0.997     0.133     0.991     0.997     0.994     0.905     0.946     0.993     negative
      0.867     0.003     0.956     0.867     0.909     0.905     0.946     0.881     sick
Weighted Avg.   0.989     0.125     0.989     0.989     0.989     0.905     0.946     0.986

==== Confusion Matrix ====

      a     b  <-- classified as
1155   3 |     a = negative
      10   65 |     b = sick

```

## Leave one out fold

```

==== Run information ====

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3627
Attributes:  30
             age
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             I131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary
             psych
             TSH_measured
             TSH
             T3_measured
             T3
             TT4_measured
             TT4
             T4U_measured
             T4U
             FTI_measured
             FTI
             TBG_measured
             TBG
             referral_source
             Class
Test mode:   3627-fold cross-validation

==== Classifier model (full training set) ====

```

```
J48 pruned tree
-----
T3 <= 1.1
|   FTI <= 155
|   |   referral_source = SVHC: sick (9.0)
|   |   referral_source = other
|   |   |   sick = f
|   |   |   |   age <= 70: negative (23.0/1.0)
|   |   |   |   age > 70
|   |   |   |   |   T4U <= 0.95: sick (11.0)
|   |   |   |   |   T4U > 0.95: negative (4.0/1.0)
|   |   |   |   sick = t: sick (3.0)
|   |   referral_source = SVI
|   |   |   TT4 <= 53
|   |   |   |   sex = F
|   |   |   |   |   age <= 55: sick (3.0)
|   |   |   |   |   age > 55: negative (10.0/1.0)
|   |   |   |   sex = M
|   |   |   |   |   age <= 51: negative (4.0/1.0)
|   |   |   |   |   age > 51: sick (4.0)
|   |   |   TT4 > 53
|   |   |   |   FTI <= 70
|   |   |   |   |   FTI <= 61: negative (2.0)
|   |   |   |   |   FTI > 61: sick (8.0)
|   |   |   |   FTI > 70: sick (153.0)
|   |   referral_source = STMW: sick (0.0)
|   |   referral_source = SVHD: sick (3.0)
FTI > 155
|   TT4 <= 92
|   |   TSH <= 0.5: negative (2.0)
|   |   TSH > 0.5: sick (3.0)
|   TT4 > 92: negative (15.0)
T3 > 1.1
|   FTI <= 171
|   TT4 <= 56
|   |   referral_source = SVHC: negative (4.0)
|   |   referral_source = other: negative (15.0/1.0)
|   |   referral_source = SVI: sick (8.0/1.0)
|   |   referral_source = STMW: negative (1.0)
|   |   referral_source = SVHD: negative (0.0)
|   TT4 > 56: negative (3230.0/1.0)
FTI > 171
|   on_thyroxine = f
|   |   T3 <= 2.7
|   |   |   T3_measured = t
|   |   |   |   TSH <= 0.31: sick (18.0/2.0)
|   |   |   |   TSH > 0.31: negative (7.0)
|   |   |   T3_measured = f: negative (15.0)
|   |   |   T3 > 2.7: negative (38.0)
|   |   on_thyroxine = t: negative (34.0)

Number of Leaves : 28

Size of the tree : 49

Time taken to build model: 0.12 seconds

==== Stratified cross-validation ====
==== Summary ====

    Correctly Classified Instances      3593           99.0626 %
    Incorrectly Classified Instances     34            0.9374 %
    Kappa statistic                      0.9194
    Mean absolute error                  0.0105
    Root mean squared error              0.093
    Relative absolute error              9.0013 %
    Root relative squared error         38.4822 %
    Total Number of Instances          3627

==== Detailed Accuracy By Class ====

          TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
          0.995     0.080     0.995     0.995     0.995     0.919     0.942     0.993     negative
          0.920     0.005     0.929     0.920     0.924     0.919     0.942     0.909     sick
Weighted Avg.   0.991     0.075     0.991     0.991     0.991     0.919     0.942     0.988

==== Confusion Matrix ====

      a      b  <-- classified as
3385  16 |      a = negative
      18  208 |      b = sick
```

### 4.5.3 Outliers

#### Cross Validation 5 folds

```

==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3548
Attributes:  30
             age
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             I131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary
             psych
             TSH_measured
             TSH
             T3_measured
             T3
             TT4_measured
             TT4
             T4U_measured
             T4U
             FTI_measured
             FTI
             TBG_measured
             TBG
             referral_source
             Class
Test mode:   5-fold cross-validation

==== Classifier model (full training set) ===

J48 pruned tree
-----
T3 <= 1.1
|  TT4 <= 53
|  |  TSH <= 13
|  |  |  FTI <= 69: sick (8.0/1.0)
|  |  |  FTI > 69: negative (7.0)
|  |  TSH > 13: negative (32.0)
|  TT4 > 53
|  |  FTI <= 155
|  |  |  referral_source = SVHC: sick (9.0)
|  |  |  referral_source = other
|  |  |  |  sick = f
|  |  |  |  |  query_hypothyroid = f
|  |  |  |  |  on_thyroxine = f
|  |  |  |  |  |  age <= 70: negative (19.0)
|  |  |  |  |  |  age > 70: sick (10.0/1.0)
|  |  |  |  |  |  on_thyroxine = t
|  |  |  |  |  |  |  FTI <= 110: sick (2.0)
|  |  |  |  |  |  |  FTI > 110: negative (4.0/1.0)
|  |  |  |  |  |  |  query_hypothyroid = t: sick (5.0/1.0)
|  |  |  |  |  |  |  sick = t: sick (3.0)
|  |  |  |  |  |  |  referral_source = SVI
|  |  |  |  |  |  |  FTI <= 70
|  |  |  |  |  |  |  |  T4U <= 1.02: sick (7.0)
|  |  |  |  |  |  |  |  T4U > 1.02: negative (2.0)
|  |  |  |  |  |  |  |  FTI > 70: sick (152.0)
|  |  |  |  |  |  |  |  referral_source = STMW: sick (0.0)
|  |  |  |  |  |  |  |  referral_source = SVHD: sick (3.0)
|  |  |  |  |  |  |  |  FTI > 155
|  |  |  |  |  |  |  |  TT4 <= 92: sick (3.0)
|  |  |  |  |  |  |  |  TT4 > 92: negative (13.0)

T3 > 1.1
|  FTI <= 171
|  |  TT4 <= 56
|  |  |  referral_source = SVHC: negative (4.0)
|  |  |  referral_source = other: negative (25.0/1.0)
|  |  |  referral_source = SVI
|  |  |  |  TSH <= 13: sick (7.0)
|  |  |  |  TSH > 13: negative (5.0)
|  |  |  |  referral_source = STMW: negative (1.0)

```

```

|   |   |   referral_source = SVHD: negative (0.0)
|   |   TT4 > 56: negative (3157.0/1.0)
|   FTI > 171
|   |   on_thyroxine = f
|   |   |   T3_measured = t
|   |   |   |   T3 <= 2.8
|   |   |   |   |   TSH <= 0.31: sick (13.0/1.0)
|   |   |   |   |   TSH > 0.31: negative (6.0)
|   |   |   |   T3 > 2.8: negative (14.0)
|   |   |   T3_measured = f: negative (12.0)
|   |   |   on_thyroxine = t: negative (25.0)

Number of Leaves : 29

Size of the tree : 51

Time taken to build model: 0.19 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3518          99.1545 %
Incorrectly Classified Instances    30           0.8455 %
Kappa statistic                   0.9257
Mean absolute error               0.0095
Root mean squared error          0.0881
Relative absolute error          8.1405 %
Root relative squared error     36.458 %
Total Number of Instances        3548

==== Detailed Accuracy By Class ====

      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC  ROC Area  PRC Area  Class
0.997    0.095    0.994    0.997    0.995    0.926    0.970    0.996    negative
0.905    0.003    0.957    0.905    0.930    0.926    0.970    0.928    sick
Weighted Avg.  0.992    0.089    0.991    0.992    0.991    0.926    0.970    0.992

==== Confusion Matrix ====

      a      b  <- classified as
3318    9 |  a = negative
  21   200 |  b = sick

```

## Cross Validation 10 folds

```

==== Run information ====

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3548
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   10-fold cross-validation

```

```

==== Classifier model (full training set) ====
J48 pruned tree
-----
T3 <= 1.1
| TT4 <= 53
| | TSH <= 13
| | | FTI <= 69: sick (8.0/1.0)
| | | FTI > 69: negative (7.0)
| | TSH > 13: negative (32.0)
| TT4 > 53
| | FTI <= 155
| | | referral_source = SVHC: sick (9.0)
| | | referral_source = other
| | | | sick = f
| | | | | query_hypothyroid = f
| | | | | on_thyroxine = f
| | | | | | age <= 70: negative (19.0)
| | | | | | age > 70: sick (10.0/1.0)
| | | | | on_thyroxine = t
| | | | | | FTI <= 110: sick (2.0)
| | | | | | FTI > 110: negative (4.0/1.0)
| | | | | query_hypothyroid = t: sick (5.0/1.0)
| | | | | sick = t: sick (3.0)
| | | | referral_source = SVI
| | | | | FTI <= 70
| | | | | | T4U <= 1.02: sick (7.0)
| | | | | | T4U > 1.02: negative (2.0)
| | | | | FTI > 70: sick (152.0)
| | | | referral_source = STMW: sick (0.0)
| | | | referral_source = SVHD: sick (3.0)
| | | FTI > 155
| | | | TT4 <= 92: sick (3.0)
| | | | TT4 > 92: negative (13.0)

T3 > 1.1
| FTI <= 171
| | TT4 <= 56
| | | referral_source = SVHC: negative (4.0)
| | | referral_source = other: negative (25.0/1.0)
| | | referral_source = SVI
| | | | TSH <= 13: sick (7.0)
| | | | TSH > 13: negative (5.0)
| | | | referral_source = STMW: negative (1.0)
| | | | referral_source = SVHD: negative (0.0)
| | | TT4 > 56: negative (3157.0/1.0)
| FTI > 171
| | on_thyroxine = f
| | | T3_measured = t
| | | | T3 <= 2.8
| | | | | TSH <= 0.31: sick (13.0/1.0)
| | | | | TSH > 0.31: negative (6.0)
| | | | | T3 > 2.8: negative (14.0)
| | | | T3_measured = f: negative (12.0)
| | | on_thyroxine = t: negative (25.0)

Number of Leaves : 29
Size of the tree : 51

Time taken to build model: 0.13 seconds

==== Stratified cross-validation ====
==== Summary ====

```

	Correctly Classified Instances	3518	99.1545 %
Incorrectly Classified Instances	30	0.8455 %	
Kappa statistic	0.927		
Mean absolute error	0.0098		
Root mean squared error	0.0896		
Relative absolute error	8.3961 %		
Root relative squared error	37.0631 %		
Total Number of Instances	3548		

```

==== Detailed Accuracy By Class ====

```

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
Weighted Avg.	0.992	0.072	0.991	0.992	0.992	0.927	0.963	0.989	
	0.996	0.077	0.995	0.996	0.995	0.927	0.963	0.995	negative
	0.923	0.004	0.940	0.923	0.932	0.927	0.963	0.910	sick

```

==== Confusion Matrix ====

```

a	b	<-- classified as
3314	13	a = negative
17	204	b = sick

## Cross Validation 15 folds

```

==== Run information ====

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3548
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   15-fold cross-validation

==== Classifier model (full training set) ====

J48 pruned tree
-----
T3 <= 1.1
|  TT4 <= 53
|  |  TSH <= 13
|  |  |  FTI <= 69: sick (8.0/1.0)
|  |  |  FTI > 69: negative (7.0)
|  |  TSH > 13: negative (32.0)
|  TT4 > 53
|  |  FTI <= 155
|  |  |  referral_source = SVHC: sick (9.0)
|  |  |  referral_source = other
|  |  |  |  sick = f
|  |  |  |  |  query_hypothyroid = f
|  |  |  |  |  on_thyroxine = f
|  |  |  |  |  |  age <= 70: negative (19.0)
|  |  |  |  |  |  age > 70: sick (10.0/1.0)
|  |  |  |  |  on_thyroxine = t
|  |  |  |  |  |  FTI <= 110: sick (2.0)
|  |  |  |  |  |  FTI > 110: negative (4.0/1.0)
|  |  |  |  |  |  query_hypothyroid = t: sick (5.0/1.0)
|  |  |  |  |  sick = t: sick (3.0)
|  |  |  |  |  referral_source = SVI
|  |  |  |  |  FTI <= 70
|  |  |  |  |  |  T4U <= 1.02: sick (7.0)
|  |  |  |  |  |  T4U > 1.02: negative (2.0)
|  |  |  |  |  |  FTI > 70: sick (152.0)
|  |  |  |  |  |  referral_source = STMW: sick (0.0)
|  |  |  |  |  |  referral_source = SVHD: sick (3.0)
|  |  |  |  |  FTI > 155
|  |  |  |  |  TT4 <= 92: sick (3.0)
|  |  |  |  |  TT4 > 92: negative (13.0)
T3 > 1.1
|  FTI <= 171
|  |  TT4 <= 56
|  |  |  referral_source = SVHC: negative (4.0)
|  |  |  referral_source = other: negative (25.0/1.0)
|  |  |  |  |  TSH <= 13: sick (7.0)
|  |  |  |  |  TSH > 13: negative (5.0)

```

```

|   |   |   referral_source = STMW: negative (1.0)
|   |   |   referral_source = SVHD: negative (0.0)
|   |   TT4 > 56: negative (3157.0/1.0)
|   FTI > 171
|   |   on_thyroxine = f
|   |   |   T3_measured = t
|   |   |   |   T3 <= 2.8
|   |   |   |   |   TSH <= 0.31: sick (13.0/1.0)
|   |   |   |   |   TSH > 0.31: negative (6.0)
|   |   |   |   T3 > 2.8: negative (14.0)
|   |   |   T3_measured = f: negative (12.0)
|   |   on_thyroxine = t: negative (25.0)

Number of Leaves : 29

Size of the tree : 51

Time taken to build model: 0.1 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3517          99.1263 %
Incorrectly Classified Instances    31           0.8737 %
Kappa statistic                   0.9244
Mean absolute error               0.0096
Root mean squared error          0.0907
Relative absolute error          8.2275 %
Root relative squared error     37.5475 %
Total Number of Instances        3548

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC     ROC Area   PRC Area   Class
      0.996     0.081     0.995     0.996     0.995     0.924     0.961     0.994     negative
      0.919     0.004     0.940     0.919     0.929     0.924     0.961     0.911     sick
Weighted Avg.      0.991     0.077     0.991     0.991     0.991     0.924     0.961     0.989

==== Confusion Matrix ====

      a     b  <- classified as
3314  13 |  a = negative
     18 203 |  b = sick

```

## Percentage Split 50%

```

==== Run information ====

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3548
Attributes:  30
             age
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             I131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary
             psych
             TSH_measured
             TSH
             T3_measured
             T3
             TT4_measured
             TT4
             T4U_measured
             T4U
             FTI_measured
             FTI
             TBG_measured
             TBG
             referral_source
             Class
Test mode:   split 50.0% train, remainder test

```

```

==== Classifier model (full training set) ====
J48 pruned tree
-----
T3 <= 1.1
| TT4 <= 53
| | TSH <= 13
| | | FTI <= 69: sick (8.0/1.0)
| | | FTI > 69: negative (7.0)
| | TSH > 13: negative (32.0)
| TT4 > 53
| | FTI <= 155
| | | referral_source = SVHC: sick (9.0)
| | | referral_source = other
| | | | sick = f
| | | | | query_hypothyroid = f
| | | | | | on_thyroxine = f
| | | | | | | age <= 70: negative (19.0)
| | | | | | | age > 70: sick (10.0/1.0)
| | | | | | on_thyroxine = t
| | | | | | | FTI <= 110: sick (2.0)
| | | | | | | FTI > 110: negative (4.0/1.0)
| | | | | | query_hypothyroid = t: sick (5.0/1.0)
| | | | | | sick = t: sick (3.0)
| | | | | | referral_source = SVI
| | | | | | | FTI <= 70
| | | | | | | T4U <= 1.02: sick (7.0)
| | | | | | | T4U > 1.02: negative (2.0)
| | | | | | FTI > 70: sick (152.0)
| | | | | | referral_source = STMW: sick (0.0)
| | | | | | referral_source = SVHD: sick (3.0)
| | | | | | FTI > 155
| | | | | | TT4 <= 92: sick (3.0)
| | | | | | TT4 > 92: negative (13.0)
T3 > 1.1
| FTI <= 171
| | TT4 <= 56
| | | referral_source = SVHC: negative (4.0)
| | | referral_source = other: negative (25.0/1.0)
| | | referral_source = SVI
| | | | TSH <= 13: sick (7.0)
| | | | TSH > 13: negative (5.0)
| | | | referral_source = STMW: negative (1.0)
| | | | referral_source = SVHD: negative (0.0)
| | | TT4 > 56: negative (3157.0/1.0)
FTI > 171
| | on_thyroxine = f
| | | T3_measured = t
| | | | T3 <= 2.8
| | | | | TSH <= 0.31: sick (13.0/1.0)
| | | | | TSH > 0.31: negative (6.0)
| | | | | T3 > 2.8: negative (14.0)
| | | | | T3_measured = f: negative (12.0)
| | | | | on_thyroxine = t: negative (25.0)

Number of Leaves : 29
Size of the tree : 51

Time taken to build model: 0.15 seconds
==== Evaluation on test split ====
Time taken to test model on test split: 0 seconds
==== Summary ====
Correctly Classified Instances      1740      98.0834 %
Incorrectly Classified Instances     34       1.9166 %
Kappa statistic                      0.8294
Mean absolute error                  0.0227
Root mean squared error              0.1342
Relative absolute error              19.3662 %
Root relative squared error         56.3653 %
Total Number of Instances           1774

==== Detailed Accuracy By Class ====
          TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
          0.990     0.168     0.989     0.990     0.990     0.829     0.896     0.987    negative
          0.832     0.010     0.848     0.832     0.840     0.829     0.896     0.761    sick
Weighted Avg.     0.981     0.159     0.981     0.981     0.981     0.829     0.896     0.973

==== Confusion Matrix ====

```

```

      a     b    <-- classified as
1651   16 |     a = negative
      18   89 |     b = sick

```

## Percentage Split 60%

```

==== Run information ====
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3548
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 60.0% train, remainder test

==== Classifier model (full training set) ====
J48 pruned tree
-----
T3 <= 1.1
| TT4 <= 53
| | TSH <= 13
| | | FTI <= 69: sick (8.0/1.0)
| | | FTI > 69: negative (7.0)
| | TSH > 13: negative (32.0)
| TT4 > 53
| | FTI <= 155
| | | referral_source = SVHC: sick (9.0)
| | | referral_source = other
| | | | sick = f
| | | | | query_hypothyroid = f
| | | | | on_thyroxine = f
| | | | | | age <= 70: negative (19.0)
| | | | | | age > 70: sick (10.0/1.0)
| | | | | on_thyroxine = t
| | | | | | FTI <= 110: sick (2.0)
| | | | | | FTI > 110: negative (4.0/1.0)
| | | | | | query_hypothyroid = t: sick (5.0/1.0)
| | | | | sick = t: sick (3.0)
| | | | | referral_source = SVI
| | | | | FTI <= 70
| | | | | | T4U <= 1.02: sick (7.0)
| | | | | | T4U > 1.02: negative (2.0)
| | | | | FTI > 70: sick (152.0)
| | | | | referral_source = STMW: sick (0.0)
| | | | | referral_source = SVHD: sick (3.0)
| | | | | FTI > 155
| | | | | | TT4 <= 92: sick (3.0)
| | | | | | TT4 > 92: negative (13.0)
T3 > 1.1
| FTI <= 171
| | TT4 <= 56
| | | referral_source = SVHC: negative (4.0)

```

```

|   |   |   referral_source = other: negative (25.0/1.0)
|   |   |   referral_source = SVI
|   |   |   |   TSH <= 13: sick (7.0)
|   |   |   |   TSH > 13: negative (5.0)
|   |   |   referral_source = STMW: negative (1.0)
|   |   |   referral_source = SVHD: negative (0.0)
|   |   TT4 > 56: negative (3157.0/1.0)
|   FTI > 171
|   on_thyroxine = f
|   |   T3_measured = t
|   |   |   T3 <= 2.8
|   |   |   |   TSH <= 0.31: sick (13.0/1.0)
|   |   |   |   TSH > 0.31: negative (6.0)
|   |   |   T3 > 2.8: negative (14.0)
|   |   |   T3_measured = f: negative (12.0)
|   |   on_thyroxine = t: negative (25.0)

Number of Leaves : 29

Size of the tree : 51

Time taken to build model: 0.14 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0.04 seconds

==== Summary ===

    Correctly Classified Instances      1399           98.5906 %
    Incorrectly Classified Instances    20            1.4094 %
    Kappa statistic                   0.866
    Mean absolute error              0.0161
    Root mean squared error          0.1157
    Relative absolute error          14.041  %
    Root relative squared error     51.5997 %
    Total Number of Instances        1419

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC   ROC Area   PRC Area   Class
      0.990     0.080     0.996     0.990     0.993     0.867     0.981     0.998   negative
      0.920     0.010     0.831     0.920     0.873     0.867     0.981     0.813   sick
Weighted Avg.   0.986     0.076     0.987     0.986     0.986     0.867     0.981     0.988

==== Confusion Matrix ===

      a      b  <- classified as
1330  14 |      a = negative
      6   69 |      b = sick

```

## Percentage Split 66%

```

==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3548
Attributes:  30
             age
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             I131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary
             psych
             TSH_measured
             TSH
             T3_measured
             T3
             TT4_measured
             TT4
             T4U_measured
             T4U

```

```

FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Test mode: split 66.0% train, remainder test

==== Classifier model (full training set) ===

J48 pruned tree
-----
T3 <= 1.1
| TT4 <= 53
| | TSH <= 13
| | | FTI <= 69: sick (8.0/1.0)
| | | FTI > 69: negative (7.0)
| | TSH > 13: negative (32.0)
| TT4 > 53
| | FTI <= 155
| | | referral_source = SVHC: sick (9.0)
| | | referral_source = other
| | | | sick = f
| | | | | query_hypothyroid = f
| | | | | | on_thyroxine = f
| | | | | | age <= 70: negative (19.0)
| | | | | | age > 70: sick (10.0/1.0)
| | | | | | on_thyroxine = t
| | | | | | FTI <= 110: sick (2.0)
| | | | | | FTI > 110: negative (4.0/1.0)
| | | | | | query_hypothyroid = t: sick (5.0/1.0)
| | | | | | sick = t: sick (3.0)
| | | | | | referral_source = SVI
| | | | | | FTI <= 70
| | | | | | | T4U <= 1.02: sick (7.0)
| | | | | | | T4U > 1.02: negative (2.0)
| | | | | | FTI > 70: sick (152.0)
| | | | | | referral_source = STMW: sick (0.0)
| | | | | | referral_source = SVHD: sick (3.0)
| | | | | | FTI > 155
| | | | | | TT4 <= 92: sick (3.0)
| | | | | | TT4 > 92: negative (13.0)

T3 > 1.1
| FTI <= 171
| | TT4 <= 56
| | | referral_source = SVHC: negative (4.0)
| | | referral_source = other: negative (25.0/1.0)
| | | referral_source = SVI
| | | | TSH <= 13: sick (7.0)
| | | | TSH > 13: negative (5.0)
| | | | referral_source = STMW: negative (1.0)
| | | | referral_source = SVHD: negative (0.0)
| | | | TT4 > 56: negative (3157.0/1.0)
| FTI > 171
| | on_thyroxine = f
| | | T3_measured = t
| | | | T3 <= 2.8
| | | | | TSH <= 0.31: sick (13.0/1.0)
| | | | | TSH > 0.31: negative (6.0)
| | | | | T3 > 2.8: negative (14.0)
| | | | | T3_measured = f: negative (12.0)
| | | | | on_thyroxine = t: negative (25.0)

Number of Leaves : 29

Size of the tree : 51

Time taken to build model: 0.1 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

==== Summary ===

Correctly Classified Instances      1185          98.2587 %
Incorrectly Classified Instances    21           1.7413 %
Kappa statistic                   0.828
Mean absolute error                0.0174
Root mean squared error            0.1264
Relative absolute error             15.5142 %
Root relative squared error       57.4788 %
Total Number of Instances         1206

==== Detailed Accuracy By Class ====

```

```

TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
0.988    0.115    0.994     0.988    0.991     0.829    0.939    0.994    negative
0.885    0.012    0.794     0.885    0.837     0.829    0.939    0.788    sick
Weighted Avg. 0.983    0.110    0.984     0.983    0.983     0.829    0.939    0.983

==== Confusion Matrix ====

a      b    <-- classified as
1131   14 |    a = negative
    7   54 |    b = sick

```

## Leave one out fold

```

==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3548
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              t131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   3548-fold cross-validation

==== Classifier model (full training set) ===

J48 pruned tree
-----
T3 <= 1.1
|  TT4 <= 53
|  |  TSH <= 13
|  |  |  FTI <= 69: sick (8.0/1.0)
|  |  |  FTI > 69: negative (7.0)
|  |  |  TSH > 13: negative (32.0)
|  TT4 > 53
|  |  FTI <= 155
|  |  |  referral_source = SVHC: sick (9.0)
|  |  |  referral_source = other
|  |  |  |  sick = f
|  |  |  |  |  query_hypothyroid = f
|  |  |  |  |  |  on_thyroxine = f
|  |  |  |  |  |  |  age <= 70: negative (19.0)
|  |  |  |  |  |  |  age > 70: sick (10.0/1.0)
|  |  |  |  |  |  on_thyroxine = t
|  |  |  |  |  |  |  FTI <= 110: sick (2.0)
|  |  |  |  |  |  |  FTI > 110: negative (4.0/1.0)
|  |  |  |  |  |  query_hypothyroid = t: sick (5.0/1.0)
|  |  |  |  |  |  sick = t: sick (3.0)
|  |  |  |  |  |  referral_source = SVI
|  |  |  |  |  |  FTI <= 70
|  |  |  |  |  |  |  T4U <= 1.02: sick (7.0)
|  |  |  |  |  |  |  T4U > 1.02: negative (2.0)
|  |  |  |  |  |  FTI > 70: sick (152.0)
|  |  |  |  |  |  referral_source = STMW: sick (0.0)
|  |  |  |  |  |  referral_source = SVHD: sick (3.0)

```

```

|   |   FTI > 155
|   |   |   TT4 <= 92: sick (3.0)
|   |   |   TT4 > 92: negative (13.0)
T3 > 1.1
|   FTI <= 171
|   |   TT4 <= 56
|   |   |   referral_source = SVHC: negative (4.0)
|   |   |   referral_source = other: negative (25.0/1.0)
|   |   |   referral_source = SVI
|   |   |   |   TSH <= 13: sick (7.0)
|   |   |   |   TSH > 13: negative (5.0)
|   |   |   referral_source = STMW: negative (1.0)
|   |   |   referral_source = SVHD: negative (0.0)
|   |   TT4 > 56: negative (3157.0/1.0)
FTI > 171
|   on_thyroxine = f
|   |   T3_measured = t
|   |   |   T3 <= 2.8
|   |   |   |   TSH <= 0.31: sick (13.0/1.0)
|   |   |   |   TSH > 0.31: negative (6.0)
|   |   |   T3 > 2.8: negative (14.0)
|   |   T3_measured = f: negative (12.0)
|   |   on_thyroxine = t: negative (25.0)

Number of Leaves : 29
Size of the tree : 51

Time taken to build model: 0.12 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3519           99.1826 %
Incorrectly Classified Instances    29            0.8174 %
Kappa statistic                   0.9293
Mean absolute error               0.0092
Root mean squared error          0.0879
Relative absolute error          7.8975 %
Root relative squared error     36.3484 %
Total Number of Instances        3548

==== Detailed Accuracy By Class ====

          TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
          0.996    0.077   0.995    0.996   0.996    0.929    0.950    0.994    negative
          0.923    0.004   0.944    0.923   0.934    0.929    0.950    0.918    sick
Weighted Avg.       0.992    0.072   0.992    0.992   0.992    0.929    0.950    0.989

==== Confusion Matrix ====
      a     b  <- classified as
3315  12 |   a = negative
      17  204 |   b = sick

```

#### 4.5.4 Conclusion

Preprocessing method	C4.5						
	Cross Validation			Percentage Split			Leave One Out Fold
	5 folds	10 folds	15 folds	50.00%	60.00%	66.00%	
With Original Dataset	98.4358%	98.8070%	98.6479%	98.3563%	98.6083%	98.0499%	98.5154%
Missing Values only	98.4624%	98.8600%	98.7540%	98.2503%	98.4758%	98.7520%	98.6744%
Missing Values + Extreme Values	98.5663%	98.8144%	99.0902%	98.3453%	98.9662%	98.9457%	99.0626%
Missing Values + Outliers	99.1545%	99.1545%	99.1263%	98.0834%	98.5906%	98.2587%	99.1826%

Just like PART, C4.5 was able to benefit from dealing with missing values in general and specially if we deal with outliers too as it won 4 tests which are all Cross validation tests and leave one out fold, while dealing with extreme values won 2 tests that are Percentage split 60% and 66%.

## 4.6 Combos

### 4.6.1 Missing Values + Outliers + Extreme Values

#### Cross Validation 5 folds

```

==== Run information ====
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3439
Attributes:  30
            age
            sex
            on_thyroxine
            query_on_thyroxine
            on_antithyroid_medication
            sick
            pregnant
            thyroid_surgery
            t131_treatment
            query_hypothyroid
            query_hyperthyroid
            lithium
            goitre
            tumor
            hypopituitary
            psych
            TSH_measured
            TSH
            T3_measured
            T3
            TT4_measured
            TT4
            T4U_measured
            T4U
            FTI_measured
            FTI
            TBG_measured
            TBG
            referral_source
            Class
Test mode:   5-fold cross-validation

==== Classifier model (full training set) ====
J48 pruned tree
-----
T3 <= 1.1
|   FTI <= 155
|   |   referral_source = SVHC: sick (9.0)
|   |   referral_source = other
|   |   |   sick = f
|   |   |   |   age <= 70: negative (21.0)
|   |   |   |   age > 70
|   |   |   |   |   T4U <= 0.95: sick (10.0)
|   |   |   |   |   T4U > 0.95: negative (4.0/1.0)
|   |   |   |   sick = t: sick (3.0)
|   |   referral_source = SVI
|   |   |   TT4 <= 49
|   |   |   |   FTI <= 64: sick (5.0)
|   |   |   |   FTI > 64: negative (7.0)
|   |   |   TT4 > 49
|   |   |   |   FTI <= 69
|   |   |   |   |   T4U <= 0.96: sick (7.0)
|   |   |   |   |   T4U > 0.96: negative (2.0)
|   |   |   |   FTI > 69: sick (154.0)
|   |   |   referral_source = STMW: sick (0.0)
|   |   |   referral_source = SVHD: sick (3.0)
FTI > 155
|   TT4 <= 92: sick (3.0)
|   TT4 > 92: negative (13.0)
T3 > 1.1
|   TT4 <= 57
|   |   referral_source = SVHC: negative (5.0)
|   |   referral_source = other: negative (15.0/1.0)
|   |   referral_source = SVI: sick (7.0)
|   |   referral_source = STMW: negative (0.0)
|   |   referral_source = SVHD: negative (0.0)
TT4 > 57
|   |   FTI <= 163: negative (3066.0)
|   |   FTI > 163
|   |   |   on_thyroxine = f
|   |   |   |   T3_measured = t

```

```

|   |   |   |   |   TT4 <= 151: negative (19.0)
|   |   |   |   |   TT4 > 151
|   |   |   |   |   T3 <= 2.8
|   |   |   |   |   |   TSH <= 0.32: sick (13.0)
|   |   |   |   |   |   TSH > 0.32: negative (6.0)
|   |   |   |   |   T3 > 2.8: negative (9.0)
|   |   |   |   T3_measured = f: negative (17.0)
|   |   |   on_thyroxine = t: negative (41.0)

Number of Leaves : 26
Size of the tree : 45

Time taken to build model: 0.11 seconds

==== Stratified cross-validation ====
==== Summary ===

Correctly Classified Instances      3411          99.1858 %
Incorrectly Classified Instances    28           0.8142 %
Kappa statistic                   0.9296
Mean absolute error               0.0097
Root mean squared error          0.0871
Relative absolute error          8.2026 %
Root relative squared error     35.9142 %
Total Number of Instances        3439

==== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC    ROC Area  PRC Area  Class
      0.997    0.083    0.994    0.997    0.996    0.930   0.978    0.997    negative
      0.917    0.003    0.952    0.917    0.934    0.930   0.978    0.928    sick
Weighted Avg.      0.992    0.078    0.992    0.992    0.992    0.930   0.978    0.993

==== Confusion Matrix ====

      a      b  <-- classified as
3213  10 |  a = negative
      18  198 |  b = sick

```

## Cross Validation 10 folds

```

==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3439
Attributes:  30
             age
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             t131_treatment
             query_hypothyroid
             query_hyperthyroid
             lithium
             goitre
             tumor
             hypopituitary
             psych
             TSH_measured
             TSH
             T3_measured
             T3
             TT4_measured
             TT4
             T4U_measured
             T4U
             FTI_measured
             FTI
             TBG_measured
             TBG
             referral_source
             Class
Test mode:   10-fold cross-validation

==== Classifier model (full training set) ====
J48 pruned tree

```

```
-----
T3 <= 1.1
|   FTI <= 155
|   |   referral_source = SVHC: sick (9.0)
|   |   referral_source = other
|   |   |   sick = f
|   |   |   |   age <= 70: negative (21.0)
|   |   |   |   age > 70
|   |   |   |   |   TAU <= 0.95: sick (10.0)
|   |   |   |   |   TAU > 0.95: negative (4.0/1.0)
|   |   |   |   sick = t: sick (3.0)
|   |   |   referral_source = SVI
|   |   |   TT4 <= 49
|   |   |   |   FTI <= 64: sick (5.0)
|   |   |   |   FTI > 64: negative (7.0)
|   |   |   TT4 > 49
|   |   |   |   FTI <= 69
|   |   |   |   |   TAU <= 0.96: sick (7.0)
|   |   |   |   |   TAU > 0.96: negative (2.0)
|   |   |   |   FTI > 69: sick (154.0)
|   |   |   referral_source = STMW: sick (0.0)
|   |   |   referral_source = SVHD: sick (3.0)
|   FTI > 155
|   |   TT4 <= 92: sick (3.0)
|   |   TT4 > 92: negative (13.0)
T3 > 1.1
|   TT4 <= 57
|   |   referral_source = SVHC: negative (5.0)
|   |   referral_source = other: negative (15.0/1.0)
|   |   referral_source = SVI: sick (7.0)
|   |   referral_source = STMW: negative (0.0)
|   |   referral_source = SVHD: negative (0.0)
|   TT4 > 57
|   |   FTI <= 163: negative (3066.0)
|   |   FTI > 163
|   |   |   on_thyroxine = f
|   |   |   |   T3_measured = t
|   |   |   |   |   TT4 <= 151: negative (19.0)
|   |   |   |   |   TT4 > 151
|   |   |   |   |   |   T3 <= 2.8
|   |   |   |   |   |   |   TSH <= 0.32: sick (13.0)
|   |   |   |   |   |   |   TSH > 0.32: negative (6.0)
|   |   |   |   |   |   T3 > 2.8: negative (9.0)
|   |   |   |   |   T3_measured = f: negative (17.0)
|   |   |   |   on_thyroxine = t: negative (41.0)

Number of Leaves : 26
Size of the tree : 45

Time taken to build model: 0.09 seconds

==== Stratified cross-validation ====
==== Summary ====

  Correctly Classified Instances      3414          99.273 %
  Incorrectly Classified Instances     25           0.727 %
  Kappa statistic                      0.9378
  Mean absolute error                  0.0082
  Root mean squared error              0.083
  Relative absolute error             6.9111 %
  Root relative squared error        34.211 %
  Total Number of Instances          3439

==== Detailed Accuracy By Class ====

          TP Rate  FP Rate  Precision  Recall   F-Measure  MCC    ROC Area  PRC Area  Class
  0.997    0.065    0.996    0.997    0.996    0.938    0.975    0.997    negative
  0.935    0.003    0.948    0.935    0.942    0.938    0.975    0.932    sick
Weighted Avg.    0.993    0.061    0.993    0.993    0.993    0.938    0.975    0.993

==== Confusion Matrix ====

  a      b  <- classified as
3212   11 |      a = negative
  14   202 |      b = sick
```

## Cross Validation 15 folds

```
==== Run information ====
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
```

```

Relation:      sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:    3439
Attributes:   30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              i131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:    15-fold cross-validation

==== Classifier model (full training set) ====
J48 pruned tree
-----
T3 <= 1.1
|   FTI <= 155
|   |   referral_source = SVHC: sick (9.0)
|   |   referral_source = other
|   |   |   sick = f
|   |   |   |   age <= 70: negative (21.0)
|   |   |   |   age > 70
|   |   |   |   |   T4U <= 0.95: sick (10.0)
|   |   |   |   |   T4U > 0.95: negative (4.0/1.0)
|   |   |   |   |   sick = t: sick (3.0)
|   |   |   |   referral_source = SVI
|   |   |   |   TT4 <= 49
|   |   |   |   |   FTI <= 64: sick (5.0)
|   |   |   |   |   FTI > 64: negative (7.0)
|   |   |   |   TT4 > 49
|   |   |   |   |   FTI <= 69
|   |   |   |   |   |   T4U <= 0.96: sick (7.0)
|   |   |   |   |   |   T4U > 0.96: negative (2.0)
|   |   |   |   |   FTI > 69: sick (154.0)
|   |   |   |   referral_source = STMW: sick (0.0)
|   |   |   |   referral_source = SVHD: sick (3.0)
|   FTI > 155
|   |   TT4 <= 92: sick (3.0)
|   |   TT4 > 92: negative (13.0)
T3 > 1.1
|   TT4 <= 57
|   |   referral_source = SVHC: negative (5.0)
|   |   referral_source = other: negative (15.0/1.0)
|   |   referral_source = SVI: sick (7.0)
|   |   referral_source = STMW: negative (0.0)
|   |   referral_source = SVHD: negative (0.0)
|   TT4 > 57
|   |   FTI <= 163: negative (3066.0)
|   |   FTI > 163
|   |   |   on_thyroxine = f
|   |   |   |   T3_measured = t
|   |   |   |   |   TT4 <= 151: negative (19.0)
|   |   |   |   |   TT4 > 151
|   |   |   |   |   |   T3 <= 2.8
|   |   |   |   |   |   |   TSH <= 0.32: sick (13.0)
|   |   |   |   |   |   |   TSH > 0.32: negative (6.0)
|   |   |   |   |   |   |   T3 > 2.8: negative (9.0)
|   |   |   |   |   T3_measured = f: negative (17.0)
|   |   |   |   on_thyroxine = t: negative (41.0)

Number of Leaves : 26

```

```

Size of the tree : 45

Time taken to build model: 0.08 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3412          99.2149 %
Incorrectly Classified Instances   27           0.7851 %
Kappa statistic                   0.9335
Mean absolute error               0.0086
Root mean squared error          0.0877
Relative absolute error          7.257  %
Root relative squared error     36.1421 %
Total Number of Instances        3439

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall    F-Measure   MCC    ROC Area   PRC Area   Class
          0.996     0.060     0.996     0.996     0.996     0.933     0.976     0.997 negative
          0.940     0.004     0.935     0.940     0.938     0.933     0.976     0.908   sick
Weighted Avg.      0.992     0.057     0.992     0.992     0.992     0.933     0.976     0.991

==== Confusion Matrix ====

      a     b  <-- classified as
3209  14 |     a = negative
      13  203 |     b = sick

```

## Percentage Split 50%

```

==== Run information ====

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3439
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 50.0% train, remainder test

==== Classifier model (full training set) ====

J48 pruned tree
-----
T3 <= 1.1
|   FTI <= 155
|   |   referral_source = SVHC: sick (9.0)
|   |   referral_source = other
|   |   |   sick = f
|   |   |   |   age <= 70: negative (21.0)
|   |   |   |   age > 70
|   |   |   |   |   T4U <= 0.95: sick (10.0)
|   |   |   |   |   T4U > 0.95: negative (4.0/1.0)

```

```

|   |   |   sick = t: sick (3.0)
|   |   referral_source = SVI
|   |   TT4 <= 49
|   |   |   FTI <= 64: sick (5.0)
|   |   |   FTI > 64: negative (7.0)
|   |   TT4 > 49
|   |   |   FTI <= 69
|   |   |   |   T4U <= 0.96: sick (7.0)
|   |   |   |   T4U > 0.96: negative (2.0)
|   |   |   FTI > 69: sick (154.0)
|   |   referral_source = STMW: sick (0.0)
|   |   referral_source = SVHD: sick (3.0)
FTI > 155
|   TT4 <= 92: sick (3.0)
|   TT4 > 92: negative (13.0)
T3 > 1.1
|   TT4 <= 57
|   |   referral_source = SVHC: negative (5.0)
|   |   referral_source = other: negative (15.0/1.0)
|   |   referral_source = SVI: sick (7.0)
|   |   referral_source = STMW: negative (0.0)
|   |   referral_source = SVHD: negative (0.0)
|   TT4 > 57
|   |   FTI <= 163: negative (3066.0)
|   |   FTI > 163
|   |   |   on_thyroxine = f
|   |   |   |   T3_measured = t
|   |   |   |   |   TT4 <= 151: negative (19.0)
|   |   |   |   |   TT4 > 151
|   |   |   |   |   |   T3 <= 2.8
|   |   |   |   |   |   |   TSH <= 0.32: sick (13.0)
|   |   |   |   |   |   |   TSH > 0.32: negative (6.0)
|   |   |   |   |   |   T3 > 2.8: negative (9.0)
|   |   |   |   |   T3_measured = f: negative (17.0)
|   |   |   on_thyroxine = t: negative (41.0)

Number of Leaves : 26
Size of the tree : 45

Time taken to build model: 0.08 seconds
==== Evaluation on test split ====
Time taken to test model on test split: 0 seconds
==== Summary ====
Correctly Classified Instances      1698          98.7784 %
Incorrectly Classified Instances     21           1.2216 %
Kappa statistic                      0.8858
Mean absolute error                  0.0176
Root mean squared error              0.1063
Relative absolute error              14.9245 %
Root relative squared error         45.5687 %
Total Number of Instances            1719

==== Detailed Accuracy By Class ====


|               | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC   | ROC Area | PRC Area | Class    |
|---------------|---------|---------|-----------|--------|-----------|-------|----------|----------|----------|
| 0.994         | 0.121   | 0.993   | 0.994     | 0.994  | 0.886     | 0.952 | 0.994    | 0.994    | negative |
| 0.879         | 0.006   | 0.906   | 0.879     | 0.892  | 0.886     | 0.952 | 0.849    | 0.849    | sick     |
| Weighted Avg. | 0.988   | 0.115   | 0.988     | 0.988  | 0.988     | 0.886 | 0.952    | 0.986    |          |


==== Confusion Matrix ====
      a      b  <- classified as
1611    9 |    a = negative
      12    87 |    b = sick

```

## Percentage Split 60%

```

==== Run information ====
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last-
Instances:   3439
Attributes:  30
             age
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication

```

```

sick
pregnant
thyroid_surgery
I131_treatment
query_hypothyroid
query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH
T3_measured
T3
TT4_measured
TT4
T4U_measured
T4U
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Test mode: split 60.0% train, remainder test

==== Classifier model (full training set) ====

J48 pruned tree
-----
T3 <= 1.1
|   FTI <= 155
|   |   referral_source = SVHC: sick (9.0)
|   |   referral_source = other
|   |   |   sick = f
|   |   |   |   age <= 70: negative (21.0)
|   |   |   |   age > 70
|   |   |   |   |   T4U <= 0.95: sick (10.0)
|   |   |   |   |   T4U > 0.95: negative (4.0/1.0)
|   |   |   |   |   sick = t: sick (3.0)
|   |   |   |   |   referral_source = SVI
|   |   |   |   |   TT4 <= 49
|   |   |   |   |   |   FTI <= 64: sick (5.0)
|   |   |   |   |   |   FTI > 64: negative (7.0)
|   |   |   |   |   TT4 > 49
|   |   |   |   |   |   FTI <= 69
|   |   |   |   |   |   T4U <= 0.96: sick (7.0)
|   |   |   |   |   |   T4U > 0.96: negative (2.0)
|   |   |   |   |   |   FTI > 69: sick (154.0)
|   |   |   |   |   |   referral_source = STMW: sick (0.0)
|   |   |   |   |   |   referral_source = SVHD: sick (3.0)
|   FTI > 155
|   |   TT4 <= 92: sick (3.0)
|   |   TT4 > 92: negative (13.0)
T3 > 1.1
|   TT4 <= 57
|   |   referral_source = SVHC: negative (5.0)
|   |   referral_source = other: negative (15.0/1.0)
|   |   referral_source = SVI: sick (7.0)
|   |   referral_source = STMW: negative (0.0)
|   |   referral_source = SVHD: negative (0.0)
|   TT4 > 57
|   |   FTI <= 163: negative (3066.0)
|   |   FTI > 163
|   |   |   on_thyroxine = f
|   |   |   |   T3_measured = t
|   |   |   |   |   TT4 <= 151: negative (19.0)
|   |   |   |   |   TT4 > 151
|   |   |   |   |   |   T3 <= 2.8
|   |   |   |   |   |   |   TSH <= 0.32: sick (13.0)
|   |   |   |   |   |   |   TSH > 0.32: negative (6.0)
|   |   |   |   |   |   |   T3 > 2.8: negative (9.0)
|   |   |   |   |   |   |   T3_measured = f: negative (17.0)
|   |   |   |   |   |   |   on_thyroxine = t: negative (41.0)

Number of Leaves : 26

Size of the tree : 45

Time taken to build model: 0.07 seconds

==== Evaluation on test split ====

Time taken to test model on test split: 0 seconds

```

```

==== Summary ====
Correctly Classified Instances      1360          98.8372 %
Incorrectly Classified Instances    16           1.1628 %
Kappa statistic                   0.9029
Mean absolute error               0.0137
Root mean squared error           0.1039
Relative absolute error           11.5296 %
Root relative squared error      42.026 %
Total Number of Instances        1376

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC     ROC Area   PRC Area   Class
      0.995     0.111     0.992     0.995     0.994     0.903     0.953     0.994     negative
      0.889     0.005     0.930     0.889     0.909     0.903     0.953     0.866     sick
Weighted Avg.       0.988     0.104     0.988     0.988     0.988     0.903     0.953     0.985

==== Confusion Matrix ====

      a      b  <-- classified as
1280    6 |  a = negative
      10    80 |  b = sick

```

## Percentage Split 66%

```

==== Run information ====
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3439
Attributes:  30
              age
              sex
              on_thyroxine
              query_on_thyroxine
              on_antithyroid_medication
              sick
              pregnant
              thyroid_surgery
              I131_treatment
              query_hypothyroid
              query_hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH_measured
              TSH
              T3_measured
              T3
              TT4_measured
              TT4
              T4U_measured
              T4U
              FTI_measured
              FTI
              TBG_measured
              TBG
              referral_source
              Class
Test mode:   split 66.0% train, remainder test

==== Classifier model (full training set) ====
J48 pruned tree
-----
T3 <= 1.1
|  FTI <= 155
|  |  referral_source = SVHC: sick (9.0)
|  |  referral_source = other
|  |  |  sick = f
|  |  |  |  age <= 70: negative (21.0)
|  |  |  |  age > 70
|  |  |  |  |  T4U <= 0.95: sick (10.0)
|  |  |  |  |  T4U > 0.95: negative (4.0/1.0)
|  |  |  |  sick = t: sick (3.0)
|  |  |  referral_source = SVI
|  |  |  |  TT4 <= 49
|  |  |  |  |  FTI <= 64: sick (5.0)
|  |  |  |  |  FTI > 64: negative (7.0)

```

```

|   |   |   TT4 > 49
|   |   |   FTI <= 69
|   |   |   |   T4U <= 0.96: sick (7.0)
|   |   |   |   T4U > 0.96: negative (2.0)
|   |   |   FTI > 69: sick (154.0)
|   |   referral_source = STMW: sick (0.0)
|   |   referral_source = SVHD: sick (3.0)
|   FTI > 155
|   |   TT4 <= 92: sick (3.0)
|   |   TT4 > 92: negative (13.0)
T3 > 1.1
|   TT4 <= 57
|   |   referral_source = SVHC: negative (5.0)
|   |   referral_source = other: negative (15.0/1.0)
|   |   referral_source = SVI: sick (7.0)
|   |   referral_source = STMW: negative (0.0)
|   |   referral_source = SVHD: negative (0.0)
|   TT4 > 57
|   |   FTI <= 163: negative (3066.0)
|   |   FTI > 163
|   |   |   on_thyroxine = f
|   |   |   T3_measured = t
|   |   |   |   TT4 <= 151: negative (19.0)
|   |   |   |   TT4 > 151
|   |   |   |   |   T3 <= 2.8
|   |   |   |   |   |   TSH <= 0.32: sick (13.0)
|   |   |   |   |   |   TSH > 0.32: negative (6.0)
|   |   |   |   |   T3 > 2.8: negative (9.0)
|   |   |   |   T3_measured = f: negative (17.0)
|   |   |   on_thyroxine = t: negative (41.0)

Number of Leaves : 26
Size of the tree : 45

Time taken to build model: 0.05 seconds
==== Evaluation on test split ====
Time taken to test model on test split: 0.01 seconds
==== Summary ====
Correctly Classified Instances      1155          98.8024 %
Incorrectly Classified Instances     14           1.1976 %
Kappa statistic                      0.8892
Mean absolute error                  0.0141
Root mean squared error              0.1059
Relative absolute error              12.1347 %
Root relative squared error         44.921  %
Total Number of Instances            1169

==== Detailed Accuracy By Class ====


|               | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC   | ROC Area | PRC Area | Class    |
|---------------|---------|---------|-----------|--------|-----------|-------|----------|----------|----------|
| negative      | 0.995   | 0.130   | 0.992     | 0.995  | 0.994     | 0.890 | 0.953    | 0.994    | negative |
| sick          | 0.870   | 0.005   | 0.923     | 0.870  | 0.896     | 0.890 | 0.953    | 0.840    | sick     |
| Weighted Avg. | 0.988   | 0.123   | 0.988     | 0.988  | 0.988     | 0.890 | 0.953    | 0.985    |          |


==== Confusion Matrix ====


| a    | b  | --> classified as |
|------|----|-------------------|
| 1095 | 5  | a = negative      |
| 9    | 60 | b = sick          |


```

## Leave one out fold

```

==== Run information ====
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3439
Attributes:  30
             age
             sex
             on_thyroxine
             query_on_thyroxine
             on_antithyroid_medication
             sick
             pregnant
             thyroid_surgery
             I131_treatment
             query_hypothyroid

```

```

query_hyperthyroid
lithium
goitre
tumor
hypopituitary
psych
TSH_measured
TSH
T3_measured
T3
TT4_measured
TT4
T4U_measured
T4U
FTI_measured
FTI
TBG_measured
TBG
referral_source
Class
Test mode: 3439-fold cross-validation

==== Classifier model (full training set) ====

J48 pruned tree
-----

T3 <= 1.1
| FTI <= 155
| | | referral_source = SVHC: sick (9.0)
| | | referral_source = other
| | | | sick = f
| | | | | age <= 70: negative (21.0)
| | | | | age > 70
| | | | | | T4U <= 0.95: sick (10.0)
| | | | | | T4U > 0.95: negative (4.0/1.0)
| | | | | sick = t: sick (3.0)
| | | | | referral_source = SVI
| | | | | TT4 <= 49
| | | | | | FTI <= 64: sick (5.0)
| | | | | | FTI > 64: negative (7.0)
| | | | | TT4 > 49
| | | | | | FTI <= 69
| | | | | | | T4U <= 0.96: sick (7.0)
| | | | | | | T4U > 0.96: negative (2.0)
| | | | | | FTI > 69: sick (154.0)
| | | | | | referral_source = STMW: sick (0.0)
| | | | | | referral_source = SVHD: sick (3.0)
| | | | | FTI > 155
| | | | | | TT4 <= 92: sick (3.0)
| | | | | | TT4 > 92: negative (13.0)

T3 > 1.1
| TT4 <= 57
| | | referral_source = SVHC: negative (5.0)
| | | referral_source = other: negative (15.0/1.0)
| | | referral_source = SVI: sick (7.0)
| | | referral_source = STMW: negative (0.0)
| | | referral_source = SVHD: negative (0.0)
| | TT4 > 57
| | | FTI <= 163: negative (3066.0)
| | | FTI > 163
| | | | on_thyroxine = f
| | | | | T3_measured = t
| | | | | | TT4 <= 151: negative (19.0)
| | | | | | TT4 > 151
| | | | | | | T3 <= 2.8
| | | | | | | | TSH <= 0.32: sick (13.0)
| | | | | | | | TSH > 0.32: negative (6.0)
| | | | | | | T3 > 2.8: negative (9.0)
| | | | | | T3_measured = f: negative (17.0)
| | | | | | on_thyroxine = t: negative (41.0)

Number of Leaves : 26
Size of the tree : 45

Time taken to build model: 0.08 seconds

==== Stratified cross-validation ====
==== Summary ===

  Correctly Classified Instances      3420          99.4475 %
  Incorrectly Classified Instances     19           0.5525 %
  Kappa statistic                      0.9524
  Mean absolute error                  0.0063
  Root mean squared error              0.0736

```

```

Relative absolute error          5.3032 %
Root relative squared error    30.3104 %
Total Number of Instances      3439

==== Detailed Accuracy By Class ====

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC     ROC Area  PRC Area  Class
0.998    0.060    0.996    0.998    0.997    0.952    0.983    0.998    negative
0.940    0.002    0.971    0.940    0.955    0.952    0.983    0.948    sick
Weighted Avg.    0.994    0.057    0.994    0.994    0.994    0.952    0.983    0.995

==== Confusion Matrix ====

  a     b  <-- classified as
3217   6 |   a = negative
  13  203 |   b = sick

```

## 4.6.2 Missing Values + Outliers + Extreme Values + frequency 40

### Cross Validation 5 folds

```

==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:    3439
Attributes:  248
              [list of attributes omitted]
Test mode:   5-fold cross-validation

==== Classifier model (full training set) ===

J48 pruned tree
-----
T3_3 = '(-inf-1.15'
|  FTI_38 = '(-inf-157.5'
|  |  referral_source = SVHC: sick (9.0)
|  |  referral_source = other
|  |  |  age_37 = '(-inf-77.5'
|  |  |  |  sick = f
|  |  |  |  |  age_33 = '(-inf-70.5)': negative (22.0)
|  |  |  |  |  age_33 = '(70.5-inf)'
|  |  |  |  |  |  T4U_13 = '(-inf-0.905)': sick (4.0)
|  |  |  |  |  |  T4U_13 = '(0.905-inf)': negative (3.0)
|  |  |  |  |  sick = t: sick (2.0)
|  |  |  |  |  age_37 = '(77.5-inf)': sick (8.0)
|  |  |  |  |  referral_source = SVI
|  |  |  |  |  TT4_1 = '(-inf-61.5'
|  |  |  |  |  |  FTI_5 = '(-inf-86.5'
|  |  |  |  |  |  |  T4U_1 = '(-inf-0.705'
|  |  |  |  |  |  |  |  FTI_3 = '(-inf-80.5)': negative (4.0/1.0)
|  |  |  |  |  |  |  |  FTI_3 = '(80.5-inf)': sick (2.0)
|  |  |  |  |  |  |  |  T4U_1 = '(0.705-inf)': sick (17.0/1.0)
|  |  |  |  |  |  |  |  FTI_5 = '(86.5-inf'
|  |  |  |  |  |  |  |  |  TSH_10 = '(-inf-0.595)': negative (3.0)
|  |  |  |  |  |  |  |  |  TSH_10 = '(0.595-inf)': sick (3.0/1.0)
|  |  |  |  |  |  |  |  |  TT4_1 = '(61.5-inf)': sick (146.0/1.0)
|  |  |  |  |  |  |  |  |  referral_source = STMW: sick (0.0)
|  |  |  |  |  |  |  |  |  referral_source = SVHD: sick (3.0)
|  |  |  |  |  |  |  |  |  FTI_38 = '(157.5-inf'
|  |  |  |  |  |  |  |  |  TT4_11 = '(-inf-92.5)': sick (3.0)
|  |  |  |  |  |  |  |  |  TT4_11 = '(92.5-inf)': negative (12.0)
T3_3 = '(1.15-inf'
|  |  |  |  |  |  |  |  |  FTI_39 = '(-inf-170.5'
|  |  |  |  |  |  |  |  |  TT4_1 = '(-inf-61.5'
|  |  |  |  |  |  |  |  |  |  age_34 = '(-inf-72.5'
|  |  |  |  |  |  |  |  |  |  |  T4U_32 = '(-inf-1.095)': negative (43.0/4.0)
|  |  |  |  |  |  |  |  |  |  |  T4U_32 = '(1.095-inf)': sick (2.0)
|  |  |  |  |  |  |  |  |  |  |  age_34 = '(72.5-inf)': sick (2.0)
|  |  |  |  |  |  |  |  |  |  |  TT4_1 = '(61.5-inf)': negative (3079.0/1.0)
|  |  |  |  |  |  |  |  |  |  |  FTI_39 = '(170.5-inf'
|  |  |  |  |  |  |  |  |  |  |  on_thyroxine = f
|  |  |  |  |  |  |  |  |  |  |  TSH_7 = '(-inf-0.295'
|  |  |  |  |  |  |  |  |  |  |  |  T3_12 = '(-inf-2.00675)': sick (7.0)
|  |  |  |  |  |  |  |  |  |  |  |  T3_12 = '(2.00675-inf'
|  |  |  |  |  |  |  |  |  |  |  |  T3_19 = '(-inf-2.65'
|  |  |  |  |  |  |  |  |  |  |  |  |  T3_measured = t
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  TSH_3 = '(-inf-0.075)': sick (5.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  TSH_3 = '(0.075-inf)': negative (2.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  T3_measured = f: negative (6.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  T3_19 = '(2.65-inf)': negative (10.0)
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  TSH_7 = '(0.295-inf)': negative (14.0)

```

```

|   |   on_thyroxine = t: negative (28.0)

Number of Leaves : 27

Size of the tree : 50

Time taken to build model: 0.23 seconds

== Stratified cross-validation ==
== Summary ==

Correctly Classified Instances      3394          98.6915 %
Incorrectly Classified Instances    45           1.3085 %
Kappa statistic                      0.8861
Mean absolute error                  0.0168
Root mean squared error              0.109
Relative absolute error              14.2245 %
Root relative squared error        44.9224 %
Total Number of Instances          3439

== Detailed Accuracy By Class ==

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC     ROC Area  PRC Area  Class
          0.995    0.130     0.991    0.995     0.993    0.886     0.973    0.996  negative
          0.870    0.005     0.917    0.870     0.893    0.886     0.973    0.855   sick
Weighted Avg.       0.987    0.122     0.987    0.987     0.987    0.886     0.973    0.987

== Confusion Matrix ==

     a     b  <-- classified as
3206  17 |     a = negative
     28 188 |     b = sick

```

## Cross Validation 10 folds

```

== Run information ==

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3439
Attributes:  248
             [list of attributes omitted]
Test mode:   10-fold cross-validation

== Classifier model (full training set) ==

J48 pruned tree
-----
T3_3 = '(-inf-1.15]'
|   FTI_38 = '(-inf-157.5]'
|   |   referral_source = SVHC: sick (9.0)
|   |   referral_source = other
|   |   |   age_37 = '(-inf-77.5]'
|   |   |   |   sick = f
|   |   |   |   |   age_33 = '(-inf-70.5)': negative (22.0)
|   |   |   |   |   age_33 = '(70.5-inf)'
|   |   |   |   |   |   T4U_13 = '(-inf-0.905)': sick (4.0)
|   |   |   |   |   |   T4U_13 = '(0.905-inf)': negative (3.0)
|   |   |   |   |   sick = t: sick (2.0)
|   |   |   age_37 = '(77.5-inf)': sick (8.0)
|   |   |   referral_source = SVI
|   |   |   |   TT4_1 = '(-inf-61.5]'
|   |   |   |   |   FTI_5 = '(-inf-86.5]'
|   |   |   |   |   T4U_1 = '(-inf-0.705]'
|   |   |   |   |   |   FTI_3 = '(-inf-80.5)': negative (4.0/1.0)
|   |   |   |   |   |   FTI_3 = '(80.5-inf)': sick (2.0)
|   |   |   |   |   T4U_1 = '(0.705-inf)': sick (17.0/1.0)
|   |   |   |   |   FTI_5 = '(86.5-inf)'
|   |   |   |   |   |   TSH_10 = '(-inf-0.595)': negative (3.0)
|   |   |   |   |   |   TSH_10 = '(0.595-inf)': sick (3.0/1.0)
|   |   |   |   TT4_1 = '(61.5-inf)': sick (146.0/1.0)
|   |   |   |   |   referral_source = STMW: sick (0.0)
|   |   |   |   |   |   referral_source = SVHD: sick (3.0)
|   |   |   |   FTI_38 = '(157.5-inf)'
|   |   |   |   |   TT4_11 = '(-inf-92.5)': sick (3.0)
|   |   |   |   |   TT4_11 = '(92.5-inf)': negative (12.0)
T3_3 = '(1.15-inf)'
|   FTI_39 = '(-inf-170.5]'
|   |   TT4_1 = '(-inf-61.5]'
|   |   |   age_34 = '(-inf-72.5]'
|   |   |   |   T4U_32 = '(-inf-1.095)': negative (43.0/4.0)
|   |   |   |   T4U_32 = '(1.095-inf)': sick (2.0)

```

```

|   |   |   age_34 = '(72.5-inf)': sick (2.0)
|   |   TT4_1 = '(61.5-inf)': negative (3079.0/1.0)
|   FTI_39 = '(170.5-inf)'
|   |   on_thyroxine = f
|   |   |   TSH_7 = '(-inf-0.295]'
|   |   |   T3_12 = '(-inf-2.00675)': sick (7.0)
|   |   |   T3_12 = '(2.00675-inf)'
|   |   |   |   T3_19 = '(-inf-2.65]'
|   |   |   |   T3_measured = t
|   |   |   |   |   TSH_3 = '(-inf-0.075)': sick (5.0)
|   |   |   |   |   TSH_3 = '(0.075-inf)': negative (2.0)
|   |   |   |   T3_measured = f: negative (6.0)
|   |   |   |   T3_19 = '(2.65-inf)': negative (10.0)
|   |   |   TSH_7 = '(0.295-inf)': negative (14.0)
|   |   on_thyroxine = t: negative (28.0)

Number of Leaves : 27

Size of the tree : 50

Time taken to build model: 0.19 seconds

== Stratified cross-validation ==
== Summary ==

Correctly Classified Instances      3394          98.6915 %
Incorrectly Classified Instances    45           1.3085 %
Kappa statistic                      0.8871
Mean absolute error                  0.0154
Root mean squared error              0.1091
Relative absolute error              13.041 %
Root relative squared error         44.9569 %
Total Number of Instances           3439

== Detailed Accuracy By Class ==

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC     ROC Area  PRC Area  Class
0.994     0.120     0.992     0.994     0.993     0.887     0.952     0.991  negative
0.880     0.006     0.909     0.880     0.894     0.887     0.952     0.843   sick
Weighted Avg.   0.987     0.113     0.987     0.987     0.987     0.887     0.952     0.982

== Confusion Matrix ==

      a      b  <-- classified as
3204   19 |     a = negative
      26  190 |     b = sick

```

## Cross Validation 15 folds

```

== Run information ==

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3439
Attributes:  248
            [list of attributes omitted]
Test mode:   15-fold cross-validation

== Classifier model (full training set) ==

J48 pruned tree
-----

T3_3 = '(-inf-1.15'
|   FTI_38 = '(-inf-157.5'
|   |   referral_source = SVHC: sick (9.0)
|   |   referral_source = other
|   |   |   age_37 = '(-inf-77.5'
|   |   |   sick = f
|   |   |   |   age_33 = '(-inf-70.5)': negative (22.0)
|   |   |   |   age_33 = '(70.5-inf)'
|   |   |   |   |   T4U_13 = '(-inf-0.905)': sick (4.0)
|   |   |   |   |   T4U_13 = '(0.905-inf)': negative (3.0)
|   |   |   |   sick = t: sick (2.0)
|   |   |   age_37 = '(77.5-inf)': sick (8.0)
|   |   |   referral_source = SVI
|   |   |   TT4_1 = '(-inf-61.5'
|   |   |   |   FTI_5 = '(-inf-86.5'
|   |   |   |   T4U_1 = '(-inf-0.705'
|   |   |   |   |   FTI_3 = '(-inf-80.5)': negative (4.0/1.0)
|   |   |   |   |   FTI_3 = '(80.5-inf)': sick (2.0)
|   |   |   |   T4U_1 = '(0.705-inf)': sick (17.0/1.0)
|   |   |   |   FTI_5 = '(86.5-inf)'
```

```

|   |   |   |   TSH_10 = '(-inf-0.595]': negative (3.0)
|   |   |   |   TSH_10 = '(0.595-inf)': sick (3.0/1.0)
|   |   |   TT4_1 = '(61.5-inf)': sick (146.0/1.0)
|   |   referral_source = STMW: sick (0.0)
|   |   referral_source = SVHD: sick (3.0)
|   FTI_38 = '(157.5-inf)'
|   |   TT4_11 = '(-inf-92.5]': sick (3.0)
|   |   TT4_11 = '(92.5-inf)': negative (12.0)
T3_3 = '(1.15-inf)'
|   FTI_39 = '(-inf-170.5]'
|   |   TT4_1 = '(-inf-61.5]'
|   |   age_34 = '(-inf-72.5]'
|   |   |   T4U_32 = '(-inf-1.095]': negative (43.0/4.0)
|   |   |   T4U_32 = '(1.095-inf)': sick (2.0)
|   |   |   age_34 = '(72.5-inf)': sick (2.0)
|   |   TT4_1 = '(61.5-inf)': negative (3079.0/1.0)
|   FTI_39 = '(170.5-inf)'
|   |   on_thyroxine = f
|   |   |   TSH_7 = '(-inf-0.295]'
|   |   |   T3_12 = '(-inf-2.00675]': sick (7.0)
|   |   |   T3_12 = '(2.00675-inf)'
|   |   |   |   T3_19 = '(-inf-2.65]'
|   |   |   |   T3_measured = t
|   |   |   |   |   TSH_3 = '(-inf-0.075]': sick (5.0)
|   |   |   |   |   TSH_3 = '(0.075-inf)': negative (2.0)
|   |   |   |   T3_measured = f: negative (6.0)
|   |   |   |   T3_19 = '(2.65-inf)': negative (10.0)
|   |   |   TSH_7 = '(0.295-inf)': negative (14.0)
|   |   on_thyroxine = t: negative (28.0)

```

Number of Leaves : 27

Size of the tree : 50

Time taken to build model: 0.19 seconds

==== Stratified cross-validation ====  
==== Summary ===

Correctly Classified Instances	3393	98.6624 %
Incorrectly Classified Instances	46	1.3376 %
Kappa statistic	0.8828	
Mean absolute error	0.0154	
Root mean squared error	0.1089	
Relative absolute error	13.0669 %	
Root relative squared error	44.8849 %	
Total Number of Instances	3439	

==== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
negative	0.995	0.139	0.991	0.995	0.993	0.883	0.956	0.991	negative
sick	0.861	0.005	0.921	0.861	0.890	0.883	0.956	0.872	sick
Weighted Avg.	0.987	0.130	0.986	0.987	0.986	0.883	0.956	0.984	

==== Confusion Matrix ===

a	b	<-- classified as
3207	16	a = negative
30	186	b = sick

## Percentage Split 50%

==== Run information ===

Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2

Relation: sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last

Instances: 3439

Attributes: 248

[list of attributes omitted]

Test mode: split 50.0% train, remainder test

==== Classifier model (full training set) ===

J48 pruned tree

```

T3_3 = '(-inf-1.15]'
|   FTI_38 = '(-inf-157.5]'
|   |   referral_source = SVHC: sick (9.0)
|   |   referral_source = other
|   |   |   age_37 = '(-inf-77.5]'
|   |   |   sick = f

```

```

|   |   |   |   age_33 = '(-inf-70.5)': negative (22.0)
|   |   |   |   age_33 = '(70.5-inf)'
|   |   |   |   T4U_13 = '(-inf-0.905)': sick (4.0)
|   |   |   |   T4U_13 = '(0.905-inf)': negative (3.0)
|   |   |   |   sick = t: sick (2.0)
|   |   |   age_37 = '(77.5-inf)': sick (8.0)
|   |   |   referral_source = SVI
|   |   |   TT4_1 = '(-inf-61.5)'
|   |   |   |   FTI_5 = '(-inf-86.5)'
|   |   |   |   T4U_1 = '(-inf-0.705)'
|   |   |   |   FTI_3 = '(-inf-80.5)': negative (4.0/1.0)
|   |   |   |   FTI_3 = '(80.5-inf)': sick (2.0)
|   |   |   |   T4U_1 = '(0.705-inf)': sick (17.0/1.0)
|   |   |   |   FTI_5 = '(86.5-inf)'
|   |   |   |   TSH_10 = '(-inf-0.595)': negative (3.0)
|   |   |   |   TSH_10 = '(0.595-inf)': sick (3.0/1.0)
|   |   |   |   TT4_1 = '(61.5-inf)': sick (146.0/1.0)
|   |   |   referral_source = STMW: sick (0.0)
|   |   |   referral_source = SVHD: sick (3.0)
|   |   FTI_38 = '(157.5-inf)'
|   |   |   TT4_11 = '(-inf-92.5)': sick (3.0)
|   |   |   TT4_11 = '(92.5-inf)': negative (12.0)
T3_3 = '(1.15-inf)'
|   |   FTI_39 = '(-inf-170.5)'
|   |   |   TT4_1 = '(-inf-61.5)'
|   |   |   age_34 = '(-inf-72.5'
|   |   |   |   T4U_32 = '(-inf-1.095)': negative (43.0/4.0)
|   |   |   |   T4U_32 = '(1.095-inf)': sick (2.0)
|   |   |   age_34 = '(72.5-inf)': sick (2.0)
|   |   |   TT4_1 = '(61.5-inf)': negative (3079.0/1.0)
FTI_39 = '(170.5-inf)'
|   |   on_thyroxine = f
|   |   TSH_7 = '(-inf-0.295'
|   |   |   T3_12 = '(-inf-2.00675)': sick (7.0)
|   |   |   T3_12 = '(2.00675-inf)'
|   |   |   |   T3_19 = '(-inf-2.65)'
|   |   |   |   T3_measured = t
|   |   |   |   TSH_3 = '(-inf-0.075)': sick (5.0)
|   |   |   |   TSH_3 = '(0.075-inf)': negative (2.0)
|   |   |   |   T3_measured = f: negative (6.0)
|   |   |   |   T3_19 = '(2.65-inf)': negative (10.0)
|   |   |   TSH_7 = '(0.295-inf)': negative (14.0)
|   |   on_thyroxine = t: negative (28.0)

Number of Leaves : 27
Size of the tree : 50

Time taken to build model: 0.26 seconds
==== Evaluation on test split ====
Time taken to test model on test split: 0.05 seconds
==== Summary ====
Correctly Classified Instances      1697          98.7202 %
Incorrectly Classified Instances    22           1.2798 %
Kappa statistic                      0.8798
Mean absolute error                  0.0198
Root mean squared error              0.1096
Relative absolute error              16.7281 %
Root relative squared error        47.0071 %
Total Number of Instances           1719

==== Detailed Accuracy By Class ====


|               | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC   | ROC Area | PRC Area | Class    |
|---------------|---------|---------|-----------|--------|-----------|-------|----------|----------|----------|
| 0.994         | 0.131   | 0.992   | 0.994     | 0.993  | 0.880     | 0.960 | 0.995    |          | negative |
| 0.869         | 0.006   | 0.905   | 0.869     | 0.887  | 0.880     | 0.960 | 0.824    |          | sick     |
| Weighted Avg. | 0.987   | 0.124   | 0.987     | 0.987  | 0.987     | 0.880 | 0.960    | 0.985    |          |


==== Confusion Matrix ====


| a    | b  | <-- classified as |
|------|----|-------------------|
| 1611 | 9  | a = negative      |
| 13   | 86 | b = sick          |


```

## Percentage Split 60%

```

==== Run information ====
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2

```

```

Relation: sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances: 3439
Attributes: 248
[list of attributes omitted]
Test mode: split 60.0% train, remainder test

==== Classifier model (full training set) ===

J48 pruned tree
-----

T3_3 = '(-inf-1.15]'
|   FTI_38 = '(-inf-157.5]'
|   |   referral_source = SVHC: sick (9.0)
|   |   referral_source = other
|   |   |   age_37 = '(-inf-77.5]'
|   |   |   |   sick = f
|   |   |   |   |   age_33 = '(-inf-70.5)': negative (22.0)
|   |   |   |   |   age_33 = '(70.5-inf)'
|   |   |   |   |   T4U_13 = '(-inf-0.905)': sick (4.0)
|   |   |   |   |   T4U_13 = '(0.905-inf)': negative (3.0)
|   |   |   |   sick = t: sick (2.0)
|   |   |   |   age_37 = '(77.5-inf)': sick (8.0)
|   |   |   |   referral_source = SVI
|   |   |   |   TT4_1 = '(-inf-61.5]'
|   |   |   |   |   FTI_5 = '(-inf-86.5]'
|   |   |   |   |   T4U_1 = '(-inf-0.705]'
|   |   |   |   |   |   FTI_3 = '(-inf-80.5)': negative (4.0/1.0)
|   |   |   |   |   |   FTI_3 = '(80.5-inf)': sick (2.0)
|   |   |   |   |   |   T4U_1 = '(0.705-inf)': sick (17.0/1.0)
|   |   |   |   |   FTI_5 = '(86.5-inf)'
|   |   |   |   |   TSH_10 = '(-inf-0.595)': negative (3.0)
|   |   |   |   |   TSH_10 = '(0.595-inf)': sick (3.0/1.0)
|   |   |   |   |   TT4_1 = '(61.5-inf)': sick (146.0/1.0)
|   |   |   |   |   referral_source = STMW: sick (0.0)
|   |   |   |   |   referral_source = SVHD: sick (3.0)
|   |   |   |   FTI_38 = '(157.5-inf'
|   |   |   |   |   TT4_11 = '(-inf-92.5)': sick (3.0)
|   |   |   |   |   TT4_11 = '(92.5-inf)': negative (12.0)
T3_3 = '(1.15-inf'
|   FTI_39 = '(-inf-170.5]'
|   |   TT4_1 = '(-inf-61.5]'
|   |   |   age_34 = '(-inf-72.5]'
|   |   |   |   T4U_32 = '(-inf-1.095)': negative (43.0/4.0)
|   |   |   |   T4U_32 = '(1.095-inf)': sick (2.0)
|   |   |   |   age_34 = '(72.5-inf)': sick (2.0)
|   |   |   TT4_1 = '(61.5-inf)': negative (3079.0/1.0)
|   |   FTI_39 = '(170.5-inf'
|   |   on_thyroxine = f
|   |   |   TSH_7 = '(-inf-0.295'
|   |   |   |   T3_12 = '(-inf-2.00675)': sick (7.0)
|   |   |   |   T3_12 = '(2.00675-inf'
|   |   |   |   |   T3_19 = '(-inf-2.65'
|   |   |   |   |   |   T3_measured = t
|   |   |   |   |   |   |   TSH_3 = '(-inf-0.075)': sick (5.0)
|   |   |   |   |   |   |   TSH_3 = '(0.075-inf)': negative (2.0)
|   |   |   |   |   |   T3_measured = f: negative (6.0)
|   |   |   |   |   |   T3_19 = '(2.65-inf)': negative (10.0)
|   |   |   |   TSH_7 = '(0.295-inf)': negative (14.0)
|   |   |   on_thyroxine = t: negative (28.0)

Number of Leaves : 27

Size of the tree : 50

Time taken to build model: 0.21 seconds

==== Evaluation on test split ===

Time taken to test model on test split: 0.03 seconds

==== Summary ===

Correctly Classified Instances      1357          98.6192 %
Incorrectly Classified Instances    19           1.3808 %
Kappa statistic                      0.8841
Mean absolute error                  0.0201
Root mean squared error              0.1151
Relative absolute error              16.9176 %
Root relative squared error         46.5369 %
Total Number of Instances            1376

==== Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area   PRC Area   Class
0.995     0.133     0.991      0.995     0.993      0.884     0.940     0.992   negative

```

```

          0.867    0.005    0.918    0.867    0.891    0.884    0.940    0.825    sick
Weighted Avg.    0.986    0.125    0.986    0.986    0.986    0.884    0.940    0.981

==== Confusion Matrix ====

      a      b  <-- classified as
1279    7 |   a = negative
     12    78 |   b = sick

```

## Percentage Split 66%

```

==== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3439
Attributes:  248
              [list of attributes omitted]
Test mode:   split 66.0% train, remainder test

==== Classifier model (full training set) ====
J48 pruned tree
-----
T3_3 = '(-inf-1.15)'
| FTI_38 = '(-inf-157.5)'
|   | referral_source = SVHC: sick (9.0)
|   | referral_source = other
|   |   age_37 = '(-inf-77.5)'
|   |   | sick = f
|   |   |   age_33 = '(-inf-70.5)': negative (22.0)
|   |   |   age_33 = '(70.5-inf)'
|   |   |   | T4U_13 = '(-inf-0.905)': sick (4.0)
|   |   |   |   T4U_13 = '(0.905-inf)': negative (3.0)
|   |   |   |   sick = t: sick (2.0)
|   |   |   age_37 = '(77.5-inf)': sick (8.0)
|   |   referral_source = SVI
|   |   | TT4_1 = '(-inf-61.5)'
|   |   |   FTI_5 = '(-inf-86.5)'
|   |   |   | T4U_1 = '(-inf-0.705)'
|   |   |   |   FTI_3 = '(-inf-80.5)': negative (4.0/1.0)
|   |   |   |   FTI_3 = '(80.5-inf)': sick (2.0)
|   |   |   |   T4U_1 = '(0.705-inf)': sick (17.0/1.0)
|   |   |   |   FTI_5 = '(86.5-inf'
|   |   |   |   TSH_10 = '(-inf-0.595)': negative (3.0)
|   |   |   |   TSH_10 = '(0.595-inf)': sick (3.0/1.0)
|   |   |   TT4_1 = '(61.5-inf)': sick (146.0/1.0)
|   |   referral_source = STMW: sick (0.0)
|   |   referral_source = SVHD: sick (3.0)
| FTI_38 = '(157.5-inf'
|   | TT4_11 = '(-inf-92.5)': sick (3.0)
|   | TT4_11 = '(92.5-inf)': negative (12.0)
T3_3 = '(1.15-inf'
| FTI_39 = '(-inf-170.5)'
|   | TT4_1 = '(-inf-61.5)'
|   |   age_34 = '(-inf-72.5'
|   |   | T4U_32 = '(-inf-1.095)': negative (43.0/4.0)
|   |   | T4U_32 = '(1.095-inf)': sick (2.0)
|   |   | age_34 = '(72.5-inf)': sick (2.0)
|   |   TT4_1 = '(61.5-inf)': negative (3079.0/1.0)
| FTI_39 = '(170.5-inf'
|   | on_thyroxine = f
|   |   TSH_7 = '(-inf-0.295'
|   |   | T3_12 = '(-inf-2.00675)': sick (7.0)
|   |   | T3_12 = '(2.00675-inf'
|   |   |   T3_19 = '(-inf-2.65'
|   |   |   | T3_measured = t
|   |   |   |   TSH_3 = '(-inf-0.075)': sick (5.0)
|   |   |   |   TSH_3 = '(0.075-inf)': negative (2.0)
|   |   |   |   T3_measured = f: negative (6.0)
|   |   |   |   T3_19 = '(2.65-inf)': negative (10.0)
|   |   |   TSH_7 = '(0.295-inf)': negative (14.0)
|   |   on_thyroxine = t: negative (28.0)

Number of Leaves : 27
Size of the tree : 50

Time taken to build model: 0.2 seconds
==== Evaluation on test split ===

```

```

Time taken to test model on test split: 0.02 seconds

==== Summary ====
Correctly Classified Instances      1153          98.6313 %
Incorrectly Classified Instances    16           1.3687 %
Kappa statistic                   0.8733
Mean absolute error               0.0202
Root mean squared error          0.1146
Relative absolute error          17.3609 %
Root relative squared error     48.623  %
Total Number of Instances        1169

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC   ROC Area   PRC Area   Class
      0.995     0.145     0.991     0.995     0.993     0.874     0.931     0.992   negative
      0.855     0.005     0.908     0.855     0.881     0.874     0.931     0.812   sick
Weighted Avg.      0.986     0.137     0.986     0.986     0.986     0.874     0.931     0.981

==== Confusion Matrix ====

      a      b  <-- classified as
1094    6 |      a = negative
      10    59 |      b = sick

```

## Leave one out fold

```

==== Run information ====
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    sick-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last
Instances:   3439
Attributes:  248
              [list of attributes omitted]
Test mode:   3439-fold cross-validation

==== Classifier model (full training set) ====
J48 pruned tree
-----
T3_3 = '(-inf-1.15)'
| FTI_38 = '(-inf-157.5)'
| | referral_source = SVHC: sick (9.0)
| | referral_source = other
| | | age_37 = '(-inf-77.5)'
| | | | sick = f
| | | | age_33 = '(-inf-70.5)': negative (22.0)
| | | | age_33 = '(70.5-inf)'
| | | | | T4U_13 = '(-inf-0.905)': sick (4.0)
| | | | | T4U_13 = '(0.905-inf)': negative (3.0)
| | | | | sick = t: sick (2.0)
| | | | age_37 = '(77.5-inf)': sick (8.0)
| | | | referral_source = SVI
| | | | TT4_1 = '(-inf-61.5)'
| | | | | FTI_5 = '(-inf-86.5'
| | | | | T4U_1 = '(-inf-0.705)'
| | | | | FTI_3 = '(-inf-80.5)': negative (4.0/1.0)
| | | | | FTI_3 = '(80.5-inf)': sick (2.0)
| | | | | T4U_1 = '(0.705-inf)': sick (17.0/1.0)
| | | | | FTI_5 = '(86.5-inf'
| | | | | TSH_10 = '(-inf-0.595)': negative (3.0)
| | | | | TSH_10 = '(0.595-inf)': sick (3.0/1.0)
| | | | | TT4_1 = '(61.5-inf)': sick (146.0/1.0)
| | | | | referral_source = STMW: sick (0.0)
| | | | | referral_source = SVHD: sick (3.0)
| | FTI_38 = '(157.5-inf'
| | | TT4_11 = '(-inf-92.5)': sick (3.0)
| | | TT4_11 = '(92.5-inf)': negative (12.0)
T3_3 = '(1.15-inf'
| FTI_39 = '(-inf-170.5'
| | TT4_1 = '(-inf-61.5'
| | | age_34 = '(-inf-72.5'
| | | | T4U_32 = '(-inf-1.095)': negative (43.0/4.0)
| | | | T4U_32 = '(1.095-inf)': sick (2.0)
| | | | age_34 = '(72.5-inf)': sick (2.0)
| | | TT4_1 = '(61.5-inf)': negative (3079.0/1.0)
| FTI_39 = '(170.5-inf'
| | on_thyroxine = f
| | | TSH_7 = '(-inf-0.295'
| | | | T3_12 = '(-inf-2.00675)': sick (7.0)
| | | | T3_12 = '(2.00675-inf'
| | | | T3_19 = '(-inf-2.65'

```

```

|   |   |   |   |   T3_measured = t
|   |   |   |   |   TSH_3 = '(-inf-0.075]': sick (5.0)
|   |   |   |   |   TSH_3 = '(0.075-inf)': negative (2.0)
|   |   |   |   T3_measured = f: negative (6.0)
|   |   |   |   T3_19 = '(2.65-inf)': negative (10.0)
|   |   |   TSH_7 = '(0.295-inf)': negative (14.0)
|   |   on_thyroxine = t: negative (28.0)

Number of Leaves : 27
Size of the tree : 50

Time taken to build model: 0.23 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      3394          98.6915 %
Incorrectly Classified Instances    45           1.3085 %
Kappa statistic                   0.8856
Mean absolute error               0.0143
Root mean squared error          0.1084
Relative absolute error          12.1339 %
Root relative squared error     44.6499 %
Total Number of Instances        3439

==== Detailed Accuracy By Class ====

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC     ROC Area   PRC Area   Class
0.995     0.134     0.991     0.995     0.993     0.886     0.947     0.993   negative
0.866     0.005     0.921     0.866     0.893     0.886     0.947     0.895   sick
Weighted Avg.                     0.987     0.126     0.987     0.987     0.987     0.886     0.947     0.987

==== Confusion Matrix ====

      a     b  <-- classified as
3207  16 |   a = negative
  29 187 |   b = sick

```

### 4.6.3 Conclusion

PARTC4.5							
Preprocessing method	Cross Validation			Percentage Split			Leave One Out Fold
	5 folds	10 folds	15 folds	50.00%	60.00%	66.00%	
With Original Dataset	98.4358%	98.8070%	98.6479%	98.3563%	98.6083%	98.0499%	98.5154%
Missing + Extreme + Outliers	99.1858%	99.2730%	99.2149%	98.7784%	98.8372%	98.8024%	99.4475%
Missing + Extreme + Outliers + frequency 40	98.6915%	98.6915%	98.6624%	98.7202%	98.6192%	98.6313%	98.6915%

The combo of dealing with missing values , extreme values and outliers helped C4.5 very much and was able to win in all tests here making, however if we add to it discretizing we see the opposite which is weird as each preprocess method by itself showed great result against the original dataset however mixing them despite increasing the performance compared to the original dataset, it wasn't able to showcase an improvement over the previous combo without discretizing.

## 4.7 General Conclusion

		C4.5										
		Preprocessing method			Cross Validation		Percentage Split		Leave One Out Fold	Avg	Order Avg	Top 3 results
		With Original Dataset			5 folds	10 folds	15 folds	50.00%	60.00%	66.00%		
Discretization	Non supervised	Equal width binning	40 bins	98.4358%	98.8070%	98.6479%	98.3363%	98.6083%	98.0499%	98.5154%	98.4887%	9
			10 bins	98.1707%	98.1972%	98.1972%	98.2503%	98.6746%	98.2839%	98.8865%	98.5846%	8
			5 bins	93.8759%	93.8759%	93.8759%	93.7964%	94.1021%	94.1498%	98.0382%	98.0606%	11
		Equal frequency binning	2 bins	93.8759%	93.8759%	93.8759%	93.7964%	94.1021%	94.1498%	93.8759%	93.9360%	15
			40 bins	98.8070%	98.8070%	98.7540%	98.7805%	98.6083%	98.5860%	98.9396%	98.8118%	3
	Supervised	With Filter	10 bins	98.1442%	98.0382%	97.7996%	98.0912%	97.9457%	97.8939%	97.8791%	97.9703%	12
			With Wrapper	98.5419%	98.4889%	98.6214%	98.4093%	98.5421%	99.0640%	98.5419%	98.6014%	7
		Missing Values only	5 bins	94.6978%	94.6713%	94.6978%	94.5387%	95.0298%	94.5398%	94.5917%	94.6810%	14
			2 bins	93.8759%	93.8759%	93.8759%	93.7964%	94.1021%	94.1498%	93.8759%	93.9360%	15
			Supervised	97.9586%	97.8526%	97.7731%	97.4549%	97.6806%	97.6599%	97.8791%	97.7513%	13
Feature Selection	With Filter	With Filter	97.9056%	98.1707%	98.1707%	98.1972%	98.1445%	97.9719%	98.1442%	98.1007%	10	
		With Wrapper	98.5419%	98.4889%	98.6214%	98.4093%	98.5421%	99.0640%	98.5419%	98.6014%	7	
	Missing Values	Missing Values + Extreme Values	98.5663%	98.8144%	99.0902%	98.3453%	98.9662%	98.9457%	99.0626%	98.8272%	2	
		Missing Values + Outliers	99.1545%	99.1545%	99.1263%	98.0834%	98.5906%	98.2587%	99.1826%	98.7929%	4	
Combos	Missing + Extreme + Outliers			99.1885%	99.2730%	99.2149%	98.7784%	98.8372%	98.8024%	99.4475%	99.0770%	1
	Missing + Extreme + Outliers + frequency 40			98.6915%	98.6915%	98.6624%	98.7202%	98.6192%	98.6313%	98.6915%	98.6725%	5

Taking a look again at the big picture of all the tests we did with C4.5 which are **119 tests** in total overall. we can see that C4.5 can actually benefit from preprocessing in general. if we look at the average scores of all the tests C4.5 with original dataset scored 9th overall which means that in general there are 8 preprocessing mechanisms that can improve the performance.

However if we want the best method it is very clear from the results that the 3 best results all came from with the combo of dealing with missing values, extreme values and outliers with **99.4475%** in leave one out fold, and **99.2730%** and **99.2149%** from Cross validation 10 and 15 folds respectively. the average score is even more impressive at **99.0770%**.

## Chapter 5

# Coding part

### 5.1 Execution with no Filters

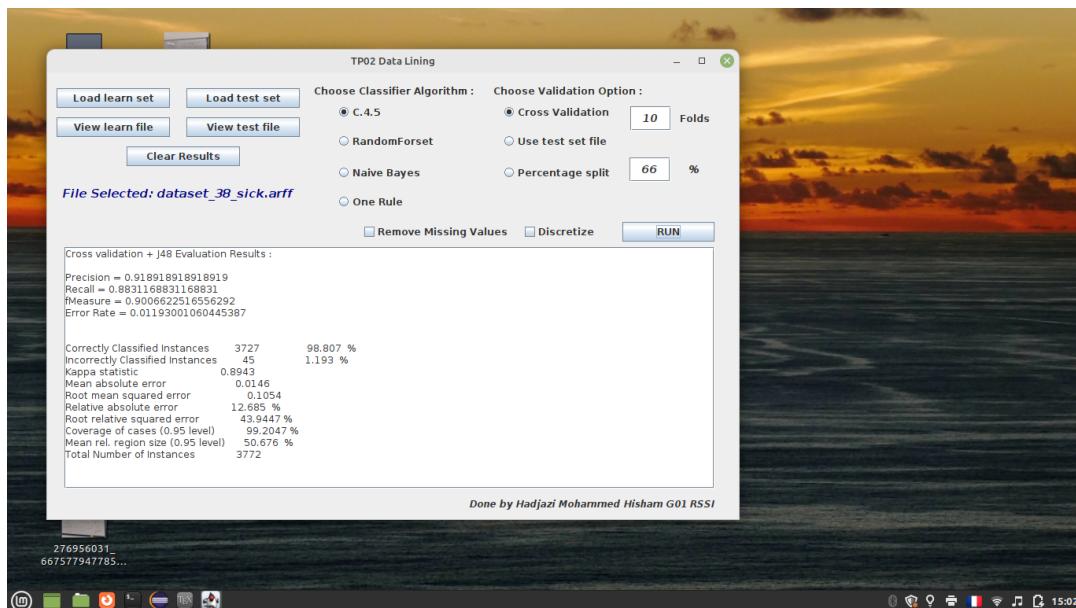


FIGURE 5.1: No Filters

### 5.2 Using Discretize() Filter

```

1 // Discretize numeric values
2 Discretize m_DiscretizeFilter = new Discretize();
3 m_DiscretizeFilter.setInputFormat(datasetInstances);
4 datasetInstances = Filter.useFilter(datasetInstances,
    m_DiscretizeFilter);

```

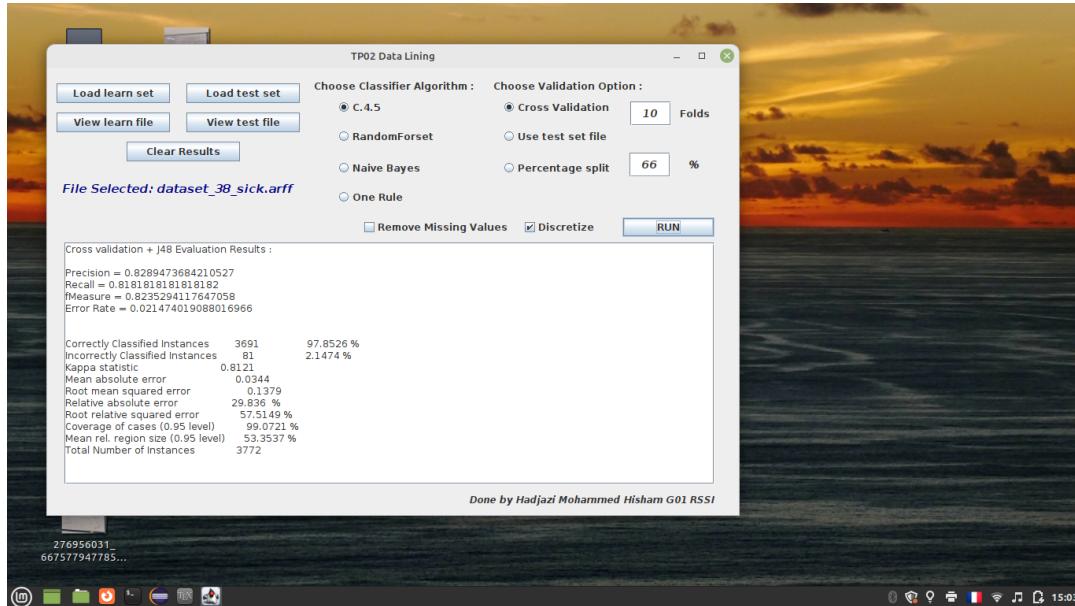


FIGURE 5.2: Discretize

### 5.3 Using ReplaceMissingValues() Filter

```

5 // remove instances with missing class
6 datasetInstances.deleteWithMissingClass();
7
8 // replace missing values
9 ReplaceMissingValues m_replaceMissing = new
10    ReplaceMissingValues();
11 m_replaceMissing.setInputFormat(datasetInstances);
12 datasetInstances = Filter.useFilter(datasetInstances,
13    m_replaceMissing);

```

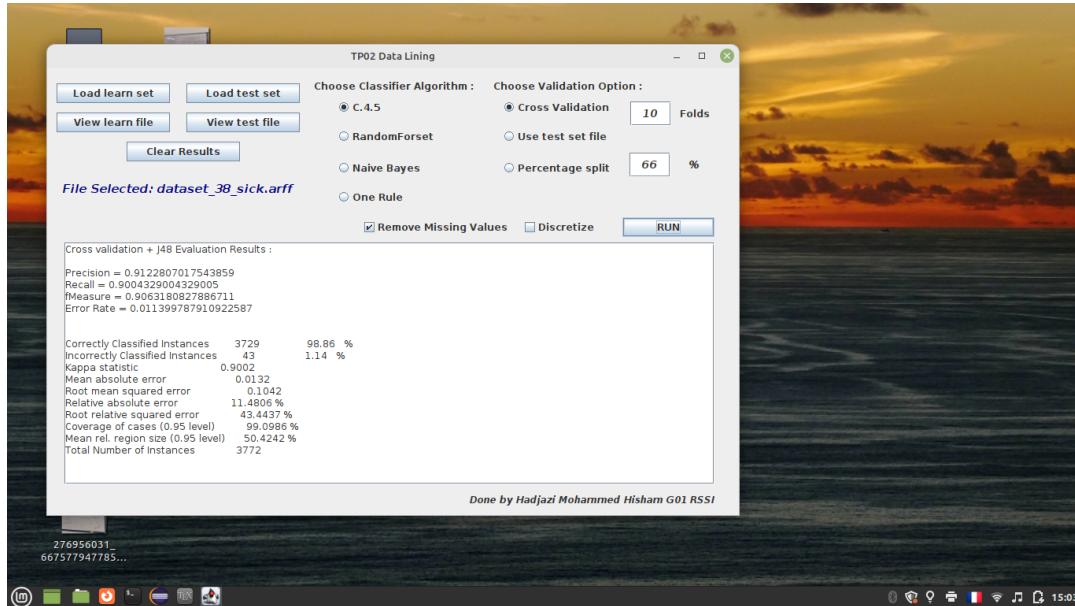


FIGURE 5.3: Replace Missing Values

## 5.4 Using ReplaceMissingValues() and Discretize() Filter

```

12 // remove instances with missing class
13 datasetInstances.deleteWithMissingClass();
14
15 // replace missing values
16 ReplaceMissingValues m_replaceMissing = new
17     ReplaceMissingValues();
18 m_replaceMissing.setInputFormat(datasetInstances);
19 datasetInstances = Filter.useFilter(datasetInstances,
20     m_replaceMissing);
21
22 // Discretize numeric values
23 Discretize m_DiscretizeFilter = new Discretize();
24 m_DiscretizeFilter.setInputFormat(datasetInstances);
25 datasetInstances = Filter.useFilter(datasetInstances,
26     m_DiscretizeFilter);

```

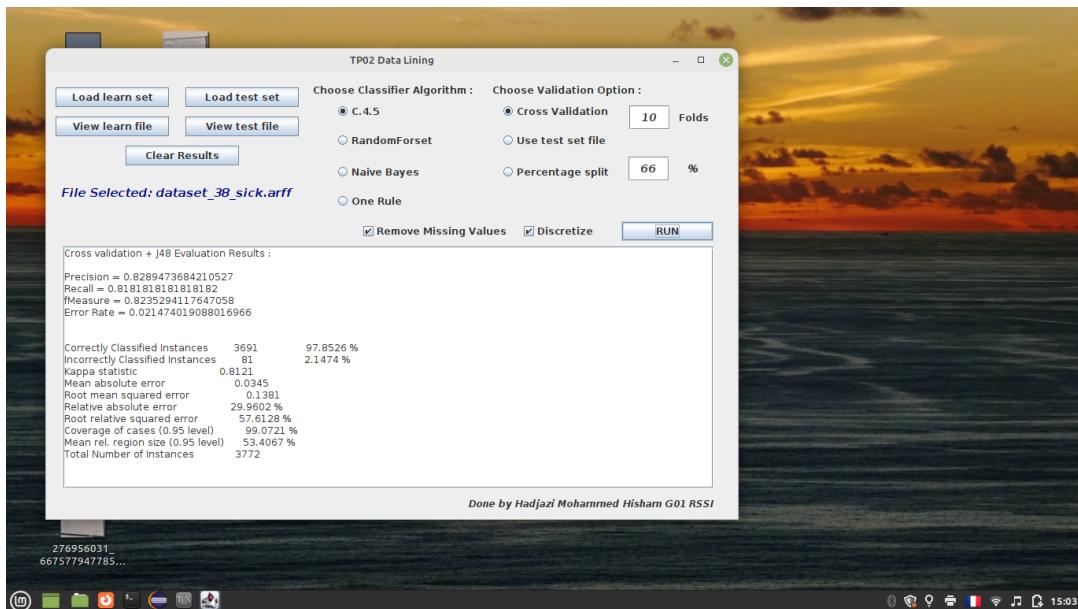


FIGURE 5.4: Replace Missing Values and Discretize

## Chapter 6

# Conclusion

This TP-03 wasn't as hard as the previous TP, the problem is in the large number of tests which is tedious work, and despite only choosing a small number of pre-processing methods 16 in total plus original dataset tests, combining them with 7 performance tests multiplies to 119 for just one classifier. and since I did only 2 classifiers as requested this means 238 total tests excluding all repeated tests due to errors or misconfiguration from my part.

Nevertheless in the first part we tested 6 classifiers against our dataset to limit our next tests on the 2 best classifiers and try to see if can improve the performance. the winners were PART and C4.5.

Starting with PART we had very hard time to find any preprocessing method to improve its performance and if see the average score of all tests we see that it is best to just stick with the original dataset without any preprocessing. however if we go to the details we can find 2 combinations that can improve our performance first is dealing with missing values with the test of Percentage split 66% we get 98.9860%. after that the second best result is given in dealing with missing , extreme and outlier values with the test of Cross Validation and 15 folds with 98.8950%.

C4.5 however did actually benefit from many preprocessing methods but one single method dominated the results which is a combo of of dealing with missing values, extreme values and outliers with 99.4475% in leave one out fold, and 99.2730% and 99.2149% from Cross validation 10 and 15 folds respectively. the average score was even more impressive at 99.0770% which actually means that C4.5 can really benefit from this preprocessing method.

In my personal tests I tried other combos and found that whenever I try mixing types of preprocessing methods i almost always get a worst result, specially when mixing with the attribute selection methods of filter and wrapper. I did try with Naive Bayes which actually was able to benefit from attribute selection. So in general the dataset and how it looks and the classifier type used can indicate from which preprocessing method it can benefit from.

Coding part this time was probably the easiest part as applying 2 filters was just adding few extra codes to my existing program, I decided to use ReplaceMissingValues() and supervised Discretize() Filters as they are probably the simplest.

## Appendix A

# Source Code

### A.1 apigui.java

```

24 package weka.api;
25
26 import java.awt.EventQueue;
27
28 import javax.swing.JFrame;
29 import javax.swing.JLabel;
30 import javax.swing.JScrollPane;
31 import javax.swing.JTextArea;
32 import javax.swing.filechooser.FileNameExtensionFilter;
33 import javax.swing.JButton;
34 import java.awt.event.ActionListener;
35 import java.io.BufferedReader;
36 import java.io.File;
37 import java.io.FileInputStream;
38 import java.io.FileReader;
39 import java.io.InputStreamReader;
40 import java.util.Random;
41 import java.awt.event.ActionEvent;
42 import javax.swing.JRadioButton;
43
44 import weka.classifiers.bayes.NaiveBayes;
45 import weka.classifiers.evaluation.Evaluation;
46 import weka.classifiers.rules.OneR;
47 import weka.classifiers.trees.J48;
48 import weka.classifiers.trees.RandomForest;
49 import weka.core.Instances;
50 import weka.filters.Filter;
51 import weka.filters.supervised.attribute.Discretize;
52 import weka.filters.unsupervised.attribute.ReplaceMissingValues;
53
54 import javax.swing.JFileChooser;
55 import javax.swing.ButtonGroup;
56 import javax.swing.SwingConstants;
57 import java.awt.Font;
58 import java.awt.SystemColor;
59 import javax.swing.JTextField;
60 import javax.swing.JCheckBox;
61
62 public class apigui {
63
64     private JFrame frmTpDataLining;
65     private File file;

```

```
66     private File file2;
67     private JRadioButton A1;
68     private JRadioButton A2;
69     private JRadioButton A3;
70     private JRadioButton A4;
71     private JRadioButton B1;
72     private JRadioButton B2;
73     private JRadioButton B3;
74     private JTextField input1;
75     private JTextField input2;
76     private int CrossValidation_Value;
77     private float Split_Percentage;
78     private BufferedReader bufferedReader;
79     private Instances datasetInstances;
80     private String DataStes;
81
82 /**
83 * Launch the application.
84 */
85 public static void main(String[] args) {
86     EventQueue.invokeLater(new Runnable() {
87         public void run() {
88             try {
89                 apigui window = new apigui();
90                 window.frmTpDataLining.setVisible(true);
91             } catch (Exception e) {
92                 e.printStackTrace();
93             }
94         }
95     });
96 }
97
98 /**
99 * Create the application.
100 */
101 public apigui() {
102     initialize();
103 }
104
105 /**
106 * Initialize the contents of the frame.
107 */
108 private void initialize() {
109
110     frmTpDataLining = new JFrame();
111     frmTpDataLining.setTitle("TP02 Data Lining");
112     frmTpDataLining.setBounds(100, 100, 881, 599);
113     frmTpDataLining.setDefaultCloseOperation(JFrame.
114         EXIT_ON_CLOSE);
115     frmTpDataLining.getContentPane().setLayout(null);
116
117     JLabel label = new JLabel("");
118     label.setForeground(SystemColor.activeCaption);
119     label.setFont(new Font("Dialog", Font.BOLD | Font.ITALIC,
120         15));
121     label.setHorizontalAlignment(SwingConstants.CENTER);
122     label.setBounds(12, 132, 309, 44);
```

```
121 frmTpDataLining.getContentPane().add(label);  
122  
123 JCheckBox ddd1 = new JCheckBox("Discretize");  
124 ddd1.setBounds(603, 191, 101, 23);  
125 frmTpDataLining.getContentPane().add(ddd1);  
126  
127 JCheckBox mmm1 = new JCheckBox("Remove Missing Values");  
128 mmm1.setBounds(399, 191, 200, 23);  
129 frmTpDataLining.getContentPane().add(mmm1);  
130  
131 JScrollPane scrollPane = new JScrollPane();  
132 scrollPane.setBounds(22, 222, 826, 307);  
133 frmTpDataLining.getContentPane().add(scrollPane);  
134  
135 JTextArea textArea = new JTextArea();  
136 scrollPane.setViewportView(textArea);  
137  
138 // Button for opening the learning dataset  
139  
140 JButton btnNewButton = new JButton("Load learn set");  
141 btnNewButton.addActionListener(new ActionListener() {  
142     @Override  
143     public void actionPerformed(ActionEvent e) {  
144         JFileChooser fileChooser = new JFileChooser();  
145         fileChooser.setFileFilter(new  
146             FileNameExtensionFilter("ARFF File", "arff"));  
147         int option = fileChooser.showOpenDialog(  
148             frmTpDataLining);  
149         if (option == JFileChooser.APPROVE_OPTION) {  
150             file = fileChooser.getSelectedFile();  
151             label.setText("File Selected: " + file.getName()  
152                ());  
153         } else {  
154             label.setText("Loading ARFF file canceled");  
155         }  
156     }  
157 } );  
158  
159 // Button for opening the learning testset  
160 JButton btnLoadTestSet = new JButton("Load test set");  
161 btnLoadTestSet.addActionListener(new ActionListener() {  
162     @Override  
163     public void actionPerformed(ActionEvent e) {  
164         JFileChooser fileChooser2 = new JFileChooser();  
165         fileChooser2.setFileFilter(new  
166             FileNameExtensionFilter("ARFF File", "arff"));  
167         int option = fileChooser2.showOpenDialog(  
168             frmTpDataLining);  
169         if (option == JFileChooser.APPROVE_OPTION) {  
170             file2 = fileChooser2.getSelectedFile();  
171             label.setText("File Selected: " + file2.getName()  
172                ());  
173         } else {  
174             label.setText("Loading TEST file canceled");  
175         }  
176     }  
177 } );
```

```
172         }
173     });
174     btnLoadTestSet.setBounds(177, 20, 144, 25);
175     frmTpDataLining.getContentPane().add(btnLoadTestSet);
176
177     btnNewButton.setBounds(12, 20, 144, 25);
178     frmTpDataLining.getContentPane().add(btnNewButton);
179
180 //    Button for view the learning learn set
181
182 JButton btnViewArff = new JButton("View learn file");
183 btnViewArff.addActionListener(new ActionListener() {
184     public void actionPerformed(ActionEvent arg0) {
185         // textArea.setText("");
186         try {
187             BufferedReader input = new BufferedReader(new
188                 InputStreamReader(new FileInputStream(file)))
189                 );
190             textArea.read(input, "READING FILE :-)");
191         } catch (Exception e) {
192             e.printStackTrace();
193         }
194     }
195 });
196 btnViewArff.setBounds(12, 57, 144, 25);
197 frmTpDataLining.getContentPane().add(btnViewArff);
198
199 input1 = new JTextField();
200 input1.setFont(new Font("Dialog", Font.BOLD | Font.ITALIC,
201     14));
202 input1.setHorizontalAlignment(SwingConstants.CENTER);
203 input1.setText("10");
204 input1.setBounds(741, 43, 52, 31);
205 frmTpDataLining.getContentPane().add(input1);
206 input1.setColumns(10);
207
208 input2 = new JTextField();
209 input2.setFont(new Font("Dialog", Font.BOLD | Font.ITALIC,
210     14));
211 input2.setText("66");
212 input2.setHorizontalAlignment(SwingConstants.CENTER);
213 input2.setColumns(10);
214 input2.setBounds(740, 108, 52, 31);
215 frmTpDataLining.getContentPane().add(input2);
216
217 JLabel lblChooseValidationOption = new JLabel("Choose
218     Classifier Algorithm :");
219 lblChooseValidationOption.setBounds(339, 16, 216, 15);
220 frmTpDataLining.getContentPane().add(
221     lblChooseValidationOption);
222
223 JLabel lblChooseValidationOption_1 = new JLabel("Choose
224     Validation Option :");
225 lblChooseValidationOption_1.setBounds(567, 16, 200, 15);
```

```
221     frmTpDataLining.getContentPane().add(  
222         lblChooseValidationOption_1);  
  
223     B1 = new JRadioButton("Cross Validation");  
224     B1.setBounds(577, 39, 150, 23);  
225     frmTpDataLining.getContentPane().add(B1);  
  
226  
227     B2 = new JRadioButton("Use test set file");  
228     B2.setBounds(577, 76, 150, 23);  
229     frmTpDataLining.getContentPane().add(B2);  
  
230  
231     B3 = new JRadioButton("Percentage split");  
232     B3.setBounds(577, 116, 150, 23);  
233     frmTpDataLining.getContentPane().add(B3);  
  
234  
235     A1 = new JRadioButton("C.4.5");  
236     A1.setBounds(367, 39, 200, 23);  
237     frmTpDataLining.getContentPane().add(A1);  
  
238  
239     A2 = new JRadioButton("RandomForset");  
240     A2.setBounds(367, 76, 200, 23);  
241     frmTpDataLining.getContentPane().add(A2);  
  
242  
243     A4 = new JRadioButton("One Rule");  
244     A4.setBounds(367, 153, 200, 23);  
245     frmTpDataLining.getContentPane().add(A4);  
  
246  
247     A3 = new JRadioButton("Naive Bayes");  
248     A3.setBounds(367, 116, 200, 23);  
249     frmTpDataLining.getContentPane().add(A3);  
  
250  
251     ButtonGroup gpp = new ButtonGroup();  
252     gpp.add(A1);  
253     gpp.add(A2);  
254     gpp.add(A3);  
255     gpp.add(A4);  
  
256  
257     ButtonGroup gpp2 = new ButtonGroup();  
258     gpp2.add(B1);  
259     gpp2.add(B2);  
260     gpp2.add(B3);  
  
261 //     Button for clearing screen  
  
262  
263     JButton btnClearResults = new JButton("Clear Results");  
264     btnClearResults.addActionListener(new ActionListener() {  
265         public void actionPerformed(ActionEvent arg0) {  
266             textArea.setText("");  
267         }  
268     });  
269     btnClearResults.setBounds(101, 94, 144, 25);  
270     frmTpDataLining.getContentPane().add(btnClearResults);  
  
271 //     Button for view the learning testset  
  
272  
273     JButton btnViewTestFile = new JButton("View test file");  
274     btnViewTestFile.addActionListener(new ActionListener() {
```

```
277     public void actionPerformed(ActionEvent arg0) {
278         // textArea.setText("");
279         try {
280             BufferedReader input = new BufferedReader(new
281                 InputStreamReader(new FileInputStream(file2)
282                     ));
283             textArea.read(input, "READING FILE :-)");
284         } catch (Exception e) {
285             e.printStackTrace();
286         }
287     });
288     btnViewTestFile.setBounds(177, 57, 144, 25);
289     frmTpDataLining.getContentPane().add(btnViewTestFile);
290
// most important button
291
292     JButton btnNewButton_2 = new JButton("RUN");
293     btnNewButton_2.addActionListener(new ActionListener() {
294         public void actionPerformed(ActionEvent arg0) {
295
296             try {
297                 // Dataset path
298                 DataStes = file.getPath();
299
300                 // Creating bufferedreader to read the dataset
301                 bufferedReader = new BufferedReader(new
302                     FileReader(DataStes));
303
304                 // Create dataset instances
305                 datasetInstances = new Instances(bufferedReader
306                     );
307
308                 if (datasetInstances.classIndex() == -1) {
309                     datasetInstances.setClassIndex(
310                         datasetInstances.numAttributes() - 1);
311                 }
312             }
313
314             // Catch block to handle the exceptions
315             catch (Exception e) {
316
317                 // Print message on the console
318                 System.out.println("Error Occurred!!!! \n" + e.
319                     getMessage());
320             }
321
322             // Missing Values + Discretize
323
324             if (ddd1.isSelected() && mmm1.isSelected()) {
325                 try {
326
327                     // remove instances with missing class
328                     datasetInstances.deleteWithMissingClass();
329
330                     // replace missing values
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2239
2240
2241
2242
2243
2244
2245
2246
2247
2248
2249
2249
2250
2251
2252
2253
2254
2255
2256
2257
2258
2259
2259
2260
2261
2262
2263
2264
2265
2266
2267
2268
2269
2269
2270
2271
2272
2273
2274
2275
2276
2277
2278
2279
2279
2280
2281
2282
2283
2284
2285
2286
2287
2288
2289
2289
2290
2291
2292
2293
2294
2295
2296
2297
2298
2299
2299
2300
2301
2302
2303
2304
2305
2306
2307
2308
2309
2309
2310
2311
2312
2313
2314
2315
2316
2317
2318
2319
2319
2320
2321
2322
2323
2324
2325
2326
2327
2328
2329
2329
2330
2331
2332
2333
2334
2335
2336
2337
2338
2339
2339
2340
2341
2342
2343
2344
2345
2346
2347
2348
2349
2349
2350
2351
2352
2353
2354
2355
2356
2357
2358
2359
2359
2360
2361
2362
2363
2364
2365
2366
2367
2368
2369
2369
2370
2371
2372
2373
2374
2375
2376
2377
2378
2379
2379
2380
2381
2382
2383
2384
2385
2386
2387
2388
2389
2389
2390
2391
2392
2393
2394
2395
2396
2397
2398
2399
2399
2400
2401
2402
2403
2404
2405
2406
2407
2408
2409
2409
2410
2411
2412
2413
2414
2415
2416
2417
2418
2419
2419
2420
2421
2422
2423
2424
2425
2426
2427
2428
2429
2429
2430
2431
2432
2433
2434
2435
2436
2437
2438
2439
2439
2440
2441
2442
2443
2444
2445
2446
2447
2448
2449
2449
2450
2451
2452
2453
2454
2455
2456
2457
2458
2459
2459
2460
2461
2462
2463
2464
2465
2466
2467
2468
2469
2469
2470
2471
2472
2473
2474
2475
2476
2477
2478
2479
2479
2480
2481
2482
2483
2484
2485
2486
2487
2488
2489
2489
2490
2491
2492
2493
2494
2495
2496
2497
2498
2499
2499
2500
2501
2502
2503
2504
2505
2506
2507
25
```

```
327     ReplaceMissingValues m_replaceMissing = new
328         ReplaceMissingValues();
329     m_replaceMissing.setInputFormat(
330         datasetInstances);
331     datasetInstances = Filter.useFilter(
332         datasetInstances, m_replaceMissing);
333
334         // Discretize numeric values
335     Discretize m_DiscretizeFilter = new
336         Discretize();
337     m_DiscretizeFilter.setInputFormat(
338         datasetInstances);
339     datasetInstances = Filter.useFilter(
340         datasetInstances, m_DiscretizeFilter);
341
342 }
343 // Catch block to handle the exceptions
344 catch (Exception e) {
345
346     // Print message on the console
347     System.out.println("Error Occurred!!!! \n"
348                     + e.getMessage());
349 }
350
351 // Discretize
352
353 } else if (ddd1.isSelected()) {
354
355     try {
356         // Discretize numeric values
357         Discretize m_DiscretizeFilter = new
358             Discretize();
359         m_DiscretizeFilter.setInputFormat(
360             datasetInstances);
361         datasetInstances = Filter.useFilter(
362             datasetInstances, m_DiscretizeFilter);
363     }
364 // Catch block to handle the exceptions
365 catch (Exception e) {
366
367     // Print message on the console
368     System.out.println("Error Occurred!!!! \n"
369                     + e.getMessage());
370 }
371
372 // Missing Values
373
374 } else if (mmml1.isSelected()) {
375
376     try {
377
378         // remove instances with missing class
379         datasetInstances.deleteWithMissingClass();
380
381         // replace missing values
382         ReplaceMissingValues m_replaceMissing = new
383             ReplaceMissingValues();
```

```
372         m_replaceMissing.setInputFormat(
373             datasetInstances);
374         datasetInstances = Filter.useFilter(
375             datasetInstances, m_replaceMissing);
376     }
377     // Catch block to handle the exceptions
378     catch (Exception e) {
379
380         // Print message on the console
381         System.out.println("Error Occurred!!!! \n"
382                         + e.getMessage());
383     }
384     // No Filters
385 } else {
386
387     try {
388         // Dataset path
389         DataStes = file.getPath();
390
391         // Creating bufferedreader to read the
392         // dataset
393         bufferedReader = new BufferedReader(new
394             FileReader(DataStes));
395
396         // Create dataset instances
397         datasetInstances = new Instances(
398             bufferedReader);
399     }
399     // Catch block to handle the exceptions
400     catch (Exception e) {
401
402         // Print message on the console
403         System.out.println("Error Occurred!!!! \n"
404                         + e.getMessage());
405     }
406 }
407
408 // Cross validation + J48
409 if (A1.isSelected() && B1.isSelected()) {
410
411     try {
412         // getting CrossValidation_Value input
413
414         CrossValidation_Value = Integer.parseInt(
415             input1.getText());
416
417         // Create J48 classifier by
418         // creating object of J48 class
419         J48 j48Classifier = new J48();
420
421         // Set Target Class
422         datasetInstances.setClassIndex(
423             datasetInstances.numAttributes() - 1);
```

```
420
421         // Evaluating by creating object of
422         // Evaluation
423         // class
424         Evaluation evaluation = new Evaluation(
425             datasetInstances);
426
427         // Cross Validate Model with user selected
428         // folds
429         evaluation.crossValidateModel(j48Classifier
430             , datasetInstances,
431             CrossValidation_Value,
432             new Random(1));
433
434         System.out.println(evaluation.
435             toSummaryString("\nResults", false));
436         textArea.setText(evaluation.toSummaryString(
437             "Cross validation + J48 Evaluation
438             Results : \n"
439             + "\nPrecision = " + evaluation.
440                 precision(1) + "\nRecall = " +
441                 evaluation.recall(1)
442             + "\nfMeasure = " + evaluation.
443                 fMeasure(1) + "\nError Rate = "
444                 + evaluation.errorRate()
445             + "\n\n", false));
446
447     }
448
449     // Catch block to handle the exceptions
450     catch (Exception e) {
451
452         // Print message on the console
453         System.out.println("Error Occurred!!!! \n"
454             + e.getMessage());
455     }
456
457 }
458 // Cross validation + RandomForest
459 else if (A2.isSelected() && B1.isSelected()) {
460
461     try {
462
463         // getting CrossValidation_Value input
464
465         CrossValidation_Value = Integer.parseInt(
466             input1.getText());
467
468         // Create RandomForest classifier by
469         // creating object of RandomForest class
470         RandomForest RandomForestClassifier = new
471             RandomForest();
472         RandomForestClassifier.setNumTrees(100);
473
474         // Set Target Class
475         datasetInstances.setClassIndex(
476             datasetInstances.numAttributes() - 1);
```



```
500     datasetInstances.setClassIndex(
501         datasetInstances.numAttributes() - 1);
502
502     // Evaluating by creating object of
503     // Evaluation
504     // class
504     Evaluation evaluation = new Evaluation(
505         datasetInstances);
506
506     // Cross Validate Model with user selected
507     // folds
507     evaluation.crossValidateModel(
508         NaiveBayesClassifier, datasetInstances,
509         CrossValidation_Value,
510         new Random(1));
511
510     System.out.println(evaluation.
511         toSummaryString("\nResults", false));
512     textArea.setText(
513         evaluation.toSummaryString("Cross
514             validation + NaiveBayes
515             Evaluation Results : \n"
516             + "\nPrecision = " +
517             evaluation.precision(1)
518             + "\nRecall = "
519             + evaluation.recall(1) + "\n
520                 nfMeasure = " +
521                 evaluation.fMeasure(1)
522             + "\nError Rate = " +
523             evaluation.errorRate() +
524             "\n\n", false));
525     }
526
527     // Catch block to handle the exceptions
528     catch (Exception e) {
529
530         // Print message on the console
531         System.out.println("Error Occurred!!!! \n"
532             + e.getMessage());
533     }
534
535     // Cross validation + One Rule
536     else if (A4.isSelected() && B1.isSelected()) {
537
538         try {
539
540             // getting CrossValidation_Value input
541
542             CrossValidation_Value = Integer.parseInt(
543                 input1.getText());
544
545             // Create One Rule classifier by
546             // creating object of One Rule class
547             OneR OneRClassifier = new OneR();
548
549         }
```

```
540 // Set Target Class
541 datasetInstances.setClassIndex(
542     datasetInstances.numAttributes() - 1);
543
544 // Evaluating by creating object of
545 // Evaluation
546 // class
547 Evaluation evaluation = new Evaluation(
548     datasetInstances);
549
550 // Cross Validate Model with user selected
551 // folds
552 evaluation.crossValidateModel(
553     OneRClassifier, datasetInstances,
554     CrossValidation_Value,
555     new Random(1));
556
557 System.out.println(evaluation.
558     toSummaryString("\nResults", false));
559 textArea.setText(
560     evaluation.toSummaryString("Cross
561     validation + One Rule Evaluation
562     Results : \n"
563     + "\nPrecision = " +
564     evaluation.precision(1)
565     + "\nRecall = "
566     + evaluation.recall(1) + "\n"
567     + "fMeasure = " +
568     evaluation.fMeasure(1)
569     + "\nError Rate = " +
570     evaluation.errorRate() +
571     "\n\n", false));
572 }
573
574 // Catch block to handle the exceptions
575 catch (Exception e) {
576
577     // Print message on the console
578     System.out.println("Error Occurred!!!! \n"
579     + e.getMessage());
580 }
581
582 }
583
584 // Test File + J48
585 else if (A1.isSelected() && B2.isSelected()) {
586
587     try {
588
589         // Create J48 classifier by
590         // creating object of J48 class
591         J48 j48Classifier = new J48();
592
593         // testset path
594         String stastes2 = file2.getPath();
595
596     }
597 }
```

```
580 // Creating bufferedreader to read the
581 // testset
582 BufferedReader bufferedReader2 = new
583 BufferedReader(new FileReader(stastes2))
584 ;
585
586 // Create test instances
587 Instances testsetInstances = new Instances(
588 bufferedReader2);
589
590 // Set Target Class
591 testsetInstances.setClassIndex(
592 testsetInstances.numAttributes() - 1);
593
594 datasetInstances.setClassIndex(
595 datasetInstances.numAttributes() - 1);
596
597 // build model
598
599 j48Classifier.buildClassifier(
600 datasetInstances);
601
602 // use
603 Evaluation eval = new Evaluation(
604 testsetInstances);
605 eval.evaluateModel(j48Classifier,
606 testsetInstances);
607
608 System.out.println(eval.toSummaryString("\
609 nResults", false));
610 textArea.setText(
611 eval.toSummaryString("Test Set +
612 J48 Evaluation Results : \n" +
613 nPrecision =
614 + eval.precision(1) + "\n
615 nRecall = " + eval.
616 recall(1) + "\nfMeasure
617 = "
618 + eval.fMeasure(1) + "\n
619 nError Rate = " + eval.
620 errorRate() + "\n\n",
621 false));
622 }
623
624 // Catch block to handle the exceptions
625 catch (Exception e) {
626
627 // Print message on the console
628 System.out.println("Error Occurred!!!! \n"
629 + e.getMessage());
630 }
631
632 }
633
634 // Test File + RandomForest
635 else if (A2.isSelected() && B2.isSelected()) {
636
637 }
```

```
618     try {
619
620         // Create RandomForest classifier by
621         // creating object of RandomForest class
622         RandomForest RandomForestClassifier = new
623             RandomForest();
624         RandomForestClassifier.setNumTrees(100);
625
626         // testset path
627         String stastes2 = file2.getPath();
628
629         // Creating bufferedreader to read the
630         // testset
631         BufferedReader bufferedReader2 = new
632             BufferedReader(new FileReader(stastes2))
633                 ;
634
635         // Create test instances
636         Instances testsetInstances = new Instances(
637             bufferedReader2);
638
639         // Set Target Class
640         testsetInstances.setClassIndex(
641             testsetInstances.numAttributes() - 1);
642
643         datasetInstances.setClassIndex(
644             datasetInstances.numAttributes() - 1);
645
646         // build model
647
648         RandomForestClassifier.buildClassifier(
649             datasetInstances);
650
651         // use
652         Evaluation eval = new Evaluation(
653             testsetInstances);
654         eval.evaluateModel(RandomForestClassifier,
655             testsetInstances);
656
657         System.out.println(eval.toSummaryString("\
658             nResults", false));
659         textArea.setText(
660             eval.toSummaryString(
661                 "Test Set + RandomForest
662                     Evaluation Results : \n"
663                     + "\nPrecision = "
664                     + eval.precision(1)
665                     + "\nRecall = "
666                     + eval.recall
667                         (1) + "\n
668                         nfMeasure = "
669                     + eval.fMeasure(1)
670                     + "\nError Rate
671                         = " + eval.
672                         errorRate() + "\n\n",
673                         false));
674     }
```

```
654     }
655
656     // Catch block to handle the exceptions
657     catch (Exception e) {
658
659         // Print message on the console
660         System.out.println("Error Occurred!!!! \n"
661                         + e.getMessage());
662     }
663 }
664
665 // Test File + NaiveBayes
666 else if (A3.isSelected() && B2.isSelected()) {
667
668     try {
669
670         // Create NaiveBayes classifier by
671         // creating object of NaiveBayes class
672         NaiveBayes NaiveBayesClassifier = new
673             NaiveBayes();
674
675         // testset path
676         String stastes2 = file2.getPath();
677
678         // Creating bufferedreader to read the
679         // testset
680         BufferedReader bufferedReader2 = new
681             BufferedReader(new FileReader(stastes2));
682
683         // Create test instances
684         Instances testsetInstances = new Instances(
685             bufferedReader2);
686
687         // Set Target Class
688         testsetInstances.setClassIndex(
689             testsetInstances.numAttributes() - 1);
690
691         datasetInstances.setClassIndex(
692             datasetInstances.numAttributes() - 1);
693
694         // build model
695
696         NaiveBayesClassifier.buildClassifier(
697             datasetInstances);
698
699         // use
700         Evaluation eval = new Evaluation(
701             testsetInstances);
702         eval.evaluateModel(NaiveBayesClassifier,
703             testsetInstances);
704
705         System.out.println(eval.toSummaryString("\
706             nResults", false));
707         textArea.setText(
708             eval.toSummaryString(
709                 "\nResults", false));
710     }
711 }
```

```
699         "Test Set + Naive Bayes  
700         Evaluation Results : \n"  
701         + "\nPrecision = "  
702         + eval.precision(1)  
703         + "\nRecall = "  
704         + eval.recall  
705         (1) + "\n  
706         nfMeasure = "  
707         + eval.fMeasure(1)  
708         + "\nError Rate  
709         = " + eval.  
710         errorRate() + "\n\n",  
711         false));  
712     }  
713  
714     // Catch block to handle the exceptions  
715     catch (Exception e) {  
716  
717         // Print message on the console  
718         System.out.println("Error Occurred!!!! \n"  
719             + e.getMessage());  
720     }  
721 }  
722  
723 // Test File + One Rule  
724 else if (A4.isSelected() && B2.isSelected()) {  
725  
726     try {  
727  
728         // Create One Rule classifier by  
729         // creating object of One Rule class  
730         OneR OneRClassifier = new OneR();  
731  
732         // testset path  
733         String stastes2 = file2.getPath();  
734  
735         // Creating bufferedreader to read the  
736         // testset  
737         BufferedReader bufferedReader2 = new  
738         BufferedReader(new FileReader(stastes2))  
739         ;  
740  
741         // Create test instances  
742         Instances testsetInstances = new Instances(  
743             bufferedReader2);  
744  
745         // Set Target Class  
746         testsetInstances.setClassIndex(  
747             testsetInstances.numAttributes() - 1);  
748  
749         datasetInstances.setClassIndex(  
750             datasetInstances.numAttributes() - 1);  
751  
752         // build model
```

```
739         OneRClassifier.buildClassifier(
    740             datasetInstances);
    741
    742         // use
    743         Evaluation eval = new Evaluation(
    744             testsetInstances);
    745         eval.evaluateModel(OneRClassifier,
    746             testsetInstances);
    747
    748         System.out.println(eval.toSummaryString("\
    749             nResults", false));
    750         textArea.setText(
    751             eval.toSummaryString(
    752                 "Test Set + One Rule
    753                     Evaluation Results : \n"
    754                     + "\nPrecision = "
    755                     + eval.precision(1)
    756                     + "\nRecall = "
    757                     + eval.recall
    758                     (1) + "\\
    759                     nfMeasure = "
    760                     + eval.fMeasure(1)
    761                     + "\nError Rate
    762                     = " + eval.
    763                     errorRate() + "\\
    764                     n\n",
    765                     false));
    766     }
    767
    768     // Catch block to handle the exceptions
    769     catch (Exception e) {
    770
    771         // Print message on the console
    772         System.out.println("Error Occurred!!!! \n"
    773             + e.getMessage());
    774     }
    775
    776     // Percentage Split + J48
    777     else if (A1.isSelected() && B3.isSelected()) {
    778
    779         try {
    780
    781             // Get Percentage from input
    782
    783             Split_Percentage = Integer.parseInt(input2.
    784                 getText());
    785
    786             // Create J48 classifier by
    787             // creating object of J48 class
    788             J48 j48Classifier = new J48();
    789
    790             // Creating bufferedreader to read the
    791             // dataset
    792             BufferedReader bufferedReader = new
    793             BufferedReader(new FileReader(DataStes));
```

```
778 //  
779 //  
780 // Create dataset instances  
781 // Instances datasetInstances = new Instances(  
782 // bufferedReader);  
783 //  
784 // String[] options = new String[4];  
785 // options[0] = "-t";  
786 // options[1] = DataStes;  
787 // options[2] = "-split-percentage";  
788 // options[3] = String.valueOf(  
// Split_Percentage);  
789 //  
790 // Randomize rand = new Randomize();  
791 // rand.setInputFormat(datasetInstances);  
792 // rand.setRandomSeed(42);  
793 // datasetInstances = Filter.useFilter(  
// datasetInstances, rand);  
794 //  
795 // get the train set  
796 // RemovePercentage rp = new RemovePercentage  
// ();  
797 // rp.setInputFormat(datasetInstances);  
798 // rp.setPercentage(Split_Percentage);  
799 // Instances test = Filter.useFilter(  
// datasetInstances, rp);  
800 //  
801 // Remove test percentage from data to get  
// the test set  
802 // rp = new RemovePercentage();  
803 // rp.setInputFormat(datasetInstances);  
804 // rp.setPercentage(Split_Percentage);  
805 // rp.setInvertSelection(true);  
806 // Instances train = Filter.useFilter(  
// datasetInstances, rp);  
807 //  
808 // int trainSize = (int) Math.round(  
809 // datasetInstances.numInstances() * Split_Percentage / 100);  
810 // int testSize = datasetInstances.  
811 // numInstances() - trainSize;  
812 // Instances train = new Instances(  
// datasetInstances, 0, trainSize);  
813 // Instances test = new Instances(  
// datasetInstances, trainSize, testSize);  
814 // Set Target Class  
815 // test.setClassIndex(test.numAttributes() -  
816 // 1);  
817 //  
818 // train.setClassIndex(train.numAttributes() -  
// 1);  
819 //  
820 // build model
```

```
821 //  
822 //  
823 //  
824 //  
825 //  
826 // System.out.println(Evaluation.  
// evaluateModel(new J48(), options));  
827 textArea.setText("Percentage Split + J48  
Evaluation Results : \n"  
+ Evaluation.evaluateModel(new J48  
( ), options));  
828 }  
829 //  
830 // Catch block to handle the exceptions  
831 catch (Exception e) {  
832 // Print message on the console  
833 System.out.println("Error Occurred!!!! \n"  
+ e.getMessage());  
834 }  
835 }  
836 //  
837 }  
838 //  
839 // Percentage Split + RandomForest  
840 else if (A2.isSelected() && B3.isSelected()) {  
841 //  
842 try {  
843 // Create RandomForest classifier by  
// creating object of RandomForest class  
844 RandomForest RandomForestClassifier = new  
RandomForest();  
RandomForestClassifier.setNumTrees(100);  
845 //  
// Creating bufferedreader to read the  
846 // dataset  
847 // BufferedReader bufferedReader = new  
BufferedReader(new FileReader(DataStes));  
848 //  
//  
// Create dataset instances  
849 Instances datasetInstances = new Instances(  
bufferedReader);  
850 //  
851 String[] options = new String[4];  
options[0] = "-t";  
options[1] = DataStes;  
options[2] = "-split-percentage";  
options[3] = String.valueOf(  
Split_Percentage);  
852 //  
// Randomize data  
853 Randomize rand = new Randomize();  
rand.setInputFormat(datasetInstances);  
rand.setRandomSeed(42);  
datasetInstances = Filter.useFilter(  
datasetInstances, rand);
```

```
868 //  
869 //           // Remove train percentage from data to  
870 //           get the train set  
871 //           RemovePercentage rp = new RemovePercentage  
872 //           ();  
873 //           Instances test = Filter.useFilter(  
874 //           datasetInstances, rp);  
875 //           // Remove test percentage from data to get  
876 //           the test set  
877 //           rp = new RemovePercentage();  
878 //           rp.setInputFormat(datasetInstances);  
879 //           rp.setPercentage(Split_Percentage);  
880 //           rp.setInvertSelection(true);  
881 //           Instances train = Filter.useFilter(  
882 //           datasetInstances, rp);  
883 //  
884 //           int trainSize = (int) Math.round(  
885 //           datasetInstances.numInstances() * Split_Percentage / 100);  
886 //           int testSize = datasetInstances.  
887 //           numInstances() - trainSize;  
888 //           Instances train = new Instances(  
889 //           datasetInstances, 0, trainSize);  
890 //           Instances test = new Instances(  
891 //           datasetInstances, trainSize, testSize);  
892 //           // Set Target Class  
893 //           test.setClassIndex(test.numAttributes() -  
894 //           1);  
895 //           train.setClassIndex(train.numAttributes() -  
896 //           1);  
897 //  
898 //           //build model  
899 //           RandomForestClassifier.buildClassifier(  
900 //           train);  
901 //           //use  
902 //           Evaluation eval = new Evaluation(test);  
903 //           eval.evaluateModel(RandomForestClassifier,  
904 //           test);  
905 //           System.out.println(eval.toSummaryString("\\  
nResults", false));  
906 //           textArea.setText(eval.toSummaryString("Test  
Set + RandomForest Evaluation Results : \\\n"+ "\nPrecision = "+  
eval.precision(1)+"\nRecall = "+eval.recall(1)+"\nFMeasure = "+  
eval.fMeasure(1)+"\nError Rate = "+eval.errorRate()+"\n\\n",  
false));
```

```
906         textArea.setText("Percentage Split + Random
907             Forset Evaluation Results : \n"
908                 + Evaluation.evaluateModel(
909                     RandomForestClassifier, options)
910             );
911     }
912
913     // Catch block to handle the exceptions
914     catch (Exception e) {
915
916         // Print message on the console
917         System.out.println("Error Occurred!!!! \n"
918             + e.getMessage());
919     }
920
921     }
922
923     // Percentage Split + NaiveBayes
924     else if (A3.isSelected() && B3.isSelected()) {
925
926         try {
927
928             // Create NaiveBayes classifier by
929             // creating object of NaiveBayes class
930             NaiveBayes NaiveBayesClassifier = new
931             NaiveBayes();
932
933             // Creating bufferedreader to read the
934             // dataset
935             BufferedReader bufferedReader = new
936             BufferedReader(new FileReader(DataStes));
937
938             // Create dataset instances
939             Instances datasetInstances = new Instances(
940             bufferedReader);
941
942             String[] options = new String[4];
943             options[0] = "-t";
944             options[1] = DataStes;
945             options[2] = "-split-percentage";
946             options[3] = String.valueOf(
947                 Split_Percentage);
948
949             // Randomize data
950             Randomize rand = new Randomize();
951             rand.setInputFormat(datasetInstances);
952             rand.setRandomSeed(42);
953             datasetInstances = Filter.useFilter(
954                 datasetInstances, rand);
955
956             // Remove train percentage from data to
957             // get the train set
958             RemovePercentage rp = new RemovePercentage
959             ();
960
961             rp.setInputFormat(datasetInstances);
962             rp.setPercentage(Split_Percentage);
```

```
951 // Instances test = Filter.useFilter(
952 // datasetInstances, rp);
953 // // Remove test percentage from data to get
954 // // the test set
955 // // rp = new RemovePercentage();
956 // // rp.setInputFormat(datasetInstances);
957 // // rp.setPercentage(Split_Percentage);
958 // // rp.setInvertSelection(true);
959 // // Instances train = Filter.useFilter(
960 // datasetInstances, rp);
961 // // int trainSize = (int) Math.round(
962 // datasetInstances.numInstances() * Split_Percentage / 100);
963 // // int testSize = datasetInstances.
964 // numInstances() - trainSize;
965 // // Instances train = new Instances(
966 // datasetInstances, 0, trainSize);
967 // // Instances test = new Instances(
968 // datasetInstances, trainSize, testSize);
969 // // // Set Target Class
970 // // test.setClassIndex(test.numAttributes() -
971 // 1);
972 // // train.setClassIndex(train.numAttributes() -
973 // 1);
974 // // //build model
975 // // NaiveBayesClassifier.buildClassifier(
976 // datasetInstances);
977 // // //use
978 // // Evaluation eval = new Evaluation(test);
979 // eval.evaluateModel(NaiveBayesClassifier,
980 // test);
981 // System.out.println(eval.toSummaryString("\
982 // nResults", false));
983 // textArea.setText(eval.toSummaryString("Test
984 // Set + Naive Bayes Evaluation Results : \n"+ "\nPrecision = "+
985 // eval.precision(1)+"\nRecall = "+eval.recall(1)+"\nFMeasure = "+
986 // eval.fMeasure(1)+"\nError Rate = "+eval.errorRate()+"\n\n",
987 // false));
988 // textArea.setText("Percentage Split + Naive
989 // Bayes Evaluation Results : \n"+
990 // + Evaluation.evaluateModel(new
991 // NaiveBayes(), options));
```

```
990
991         // Print message on the console
992         System.out.println("Error Occurred!!!! \n"
993                         + e.getMessage());
994     }
995 }
996
997 // Percentage Split + One Rule
998 else if (A4.isSelected() && B3.isSelected()) {
999
1000    try {
1001
1002        // Create One Rule classifier by
1003        // creating object of One Rule class
1004        OneR OneRClassifier = new OneR();
1005
1006        // Creating bufferedreader to read the
1007        // dataset
1008        // BufferedReader bufferedReader = new
1009        // BufferedReader(new FileReader(DataStes));
1010
1011        // Create dataset instances
1012        Instances datasetInstances = new Instances(
1013            bufferedReader);
1014
1015        String[] options = new String[4];
1016        options[0] = "-t";
1017        options[1] = DataStes;
1018        options[2] = "-split-percentage";
1019        options[3] = String.valueOf(
1020            Split_Percentage);
1021
1022        // Randomize data
1023        // Randomize rand = new Randomize();
1024        // rand.setInputFormat(datasetInstances);
1025        // rand.setRandomSeed(42);
1026        // datasetInstances = Filter.useFilter(
1027        //     datasetInstances, rand);
1028
1029        // Remove train percentage from data to
1030        // get the train set
1031        // RemovePercentage rp = new RemovePercentage
1032        // ();
1033
1034        // rp.setInputFormat(datasetInstances);
1035        // rp.setPercentage(Split_Percentage);
1036        // Instances test = Filter.useFilter(
1037        //     datasetInstances, rp);
1038
1039        // Remove test percentage from data to get
1040        // the test set
1041        // rp = new RemovePercentage();
1042        // rp.setInputFormat(datasetInstances);
1043        // rp.setPercentage(Split_Percentage);
1044        // rp.setInvertSelection(true);
```

```
1036 // Instances train = Filter.useFilter(
1037 // datasetInstances, rp);
1038 //
1039 //
1040 // int trainSize = (int) Math.round(
1041 // datasetInstances.numInstances() * Split_Percentage / 100);
1042 // int testSize = datasetInstances.
1043 numInstances() - trainSize;
1044 // Instances train = new Instances(
1045 datasetInstances, 0, trainSize);
1046 // Instances test = new Instances(
1047 datasetInstances, trainSize, testSize);
1048 //
1049 // // Set Target Class
1050 // test.setClassIndex(test.numAttributes() -
1051 // 1);
1052 // //
1053 // // build model
1054 // OneRClassifier.buildClassifier(train);
1055 // //
1056 // // use
1057 // Evaluation eval = new Evaluation(test);
1058 // //
1059 // eval.evaluateModel(OneRClassifier,test);
1060 // System.out.println(eval.toSummaryString("\
1061 // nResults", false));
1062 // textArea.setText(eval.toSummaryString("Test
1063 Set + One Rule Evaluation Results : \n"+ "\nPrecision = "+eval.
precision(1)+"\nRecall = "+eval.recall(1)+"\nfMeasure = "+eval.
fMeasure(1)+"\nError Rate = "+eval.errorRate()+"\n\n", false));
1064 textArea.setText("Percentage Split + One
1065 Rule Evaluation Results : \n"
1066 + Evaluation.evaluateModel(new OneR
1067 (), options));
1068 }
1069 //
1070 // Catch block to handle the exceptions
1071 catch (Exception e) {
1072 //
1073 // Print message on the console
1074 System.out.println("Error Occurred!!!! \n"
1075 + e.getMessage());
1076 }
1077 }
1078 });
1079 btnNewButton_2.setBounds(731, 190, 117, 25);
1080 frmTpDataLining.getContentPane().add(btnNewButton_2);
```

```
1078
1079     JLabel lblNewLabel = new JLabel("Done by Hadjazi Mohammed
1080         Hisham G01 RSSI");
1081     lblNewLabel.setFont(new Font("Dialog", Font.BOLD | Font.
1082             ITALIC, 12));
1083     lblNewLabel.setHorizontalAlignment(SwingConstants.RIGHT);
1084     lblNewLabel.setBounds(494, 541, 354, 15);
1085     frmTpDataLining.getContentPane().add(lblNewLabel);
1086
1087     JLabel lblFolds = new JLabel("Folds");
1088     lblFolds.setHorizontalAlignment(SwingConstants.CENTER);
1089     lblFolds.setBounds(797, 51, 52, 15);
1090     frmTpDataLining.getContentPane().add(lblFolds);
1091
1092     JLabel lblFolds_1 = new JLabel("%");
1093     lblFolds_1.setFont(new Font("Dialog", Font.BOLD | Font.
1094             ITALIC, 14));
1095     lblFolds_1.setHorizontalAlignment(SwingConstants.CENTER);
1096     lblFolds_1.setBounds(796, 116, 52, 15);
1097     frmTpDataLining.getContentPane().add(lblFolds_1);
1098
1099 }
```

# Bibliography

- [1] *7 Types of Classification Algorithms*. en-US. Jan. 2018. URL: <https://analyticsindiamag.com/7-types-classification-algorithms/> (visited on 03/05/2022).
- [2] Shawkat Ali and Kate A. Smith. "On learning algorithm selection for classification". en. In: *Applied Soft Computing* 6.2 (Jan. 2006), pp. 119–138. ISSN: 15684946. DOI: 10.1016/j.asoc.2004.12.002. URL: <https://linkinghub.elsevier.com/retrieve/pii/S1568494605000049> (visited on 03/05/2022).
- [3] *Data Mining Map*. URL: <https://www.saedsayad.com/> (visited on 03/05/2022).
- [4] *Data mining: the textbook*. 1st edition. New York, NY: Springer Science+Business Media, 2015. ISBN: 9783319141411.
- [5] *Decision Trees – C4.5*. en. Mar. 2011. URL: <https://octaviansima.wordpress.com/2011/03/25/decision-trees-c4-5/> (visited on 03/05/2022).
- [6] *Documentation - Weka Wiki*. URL: <https://waikato.github.io/weka-wiki/documentation/> (visited on 03/06/2022).
- [7] Uday Kamath and Krishna Choppella. *Mastering Java machine learning: a java developer's guide to implementing machine learning and big data architectures*. English. OCLC: 1059420113. 2017. ISBN: 9781785880513.
- [8] *UCI Machine Learning Repository: Thyroid Disease Data Set*. URL: <http://archive.ics.uci.edu/ml/datasets/thyroid+disease> (visited on 03/27/2022).
- [9] Joaquin Vanschoren. *OpenML*. en. URL: <https://www.openml.org> (visited on 03/27/2022).
- [10] I. H. Witten, Eibe Frank, and Mark A. Hall. *Data mining: practical machine learning tools and techniques*. 3rd ed. Morgan Kaufmann series in data management systems. OCLC: ocn262433473. Burlington, MA: Morgan Kaufmann, 2011. ISBN: 9780123748560.
- [11] I. H. Witten and I. H. Witten, eds. *Data mining: practical machine learning tools and techniques*. Fourth Edition. Amsterdam: Elsevier, 2017. ISBN: 9780128042915.
- [12] Tony Yiu. *Understanding Random Forest*. en. Sept. 2021. URL: <https://towardsdatascience.com/understanding-random-forest-58381e0602d2> (visited on 03/05/2022).