



پیاده سازی پروژه درس مدل سازی و شناسایی الگو با داده های تومر در بدن

استاد راهنما:

آقای دکتر محمد حسین معطر

ارائه دهندگان:

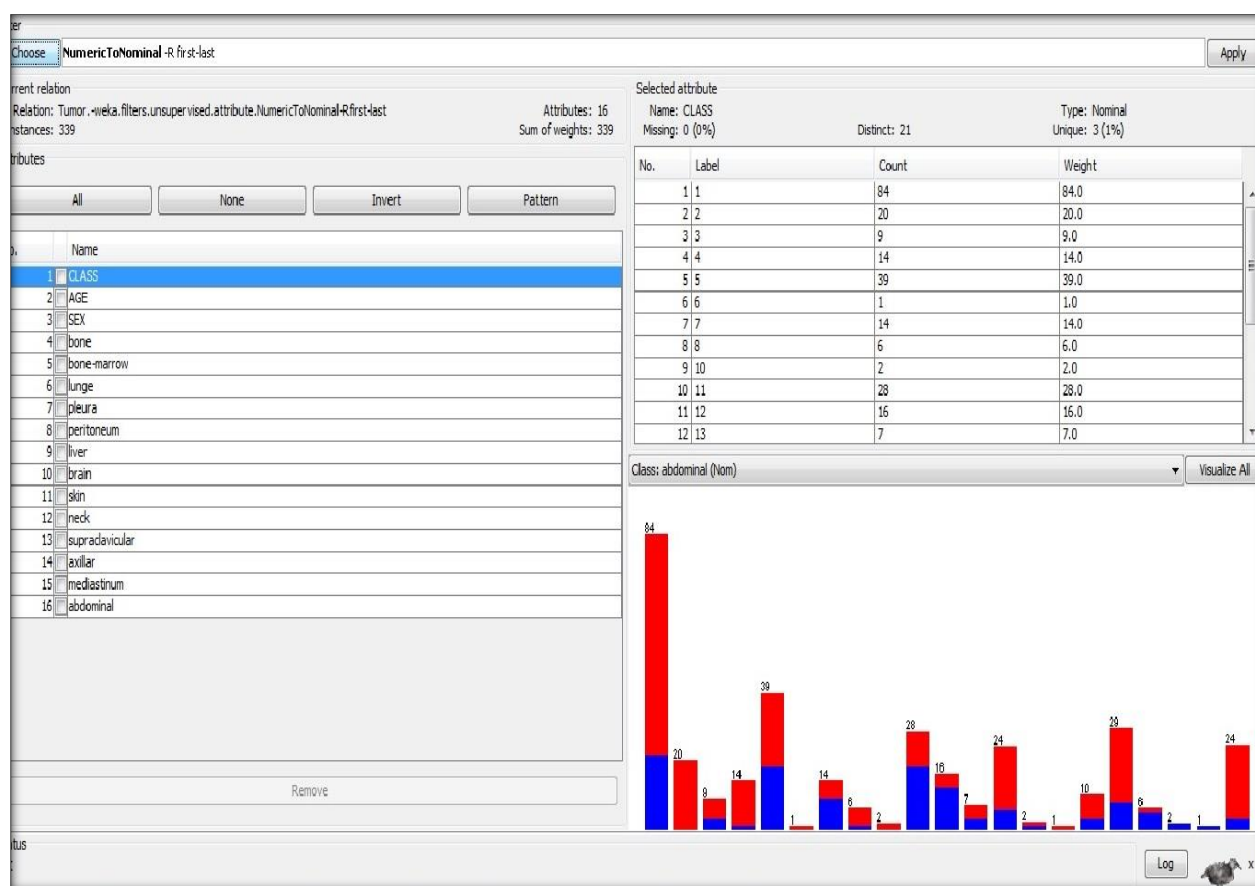
فاطمه کاکائی-سمیراضیایی

پیاده سازی پروژه

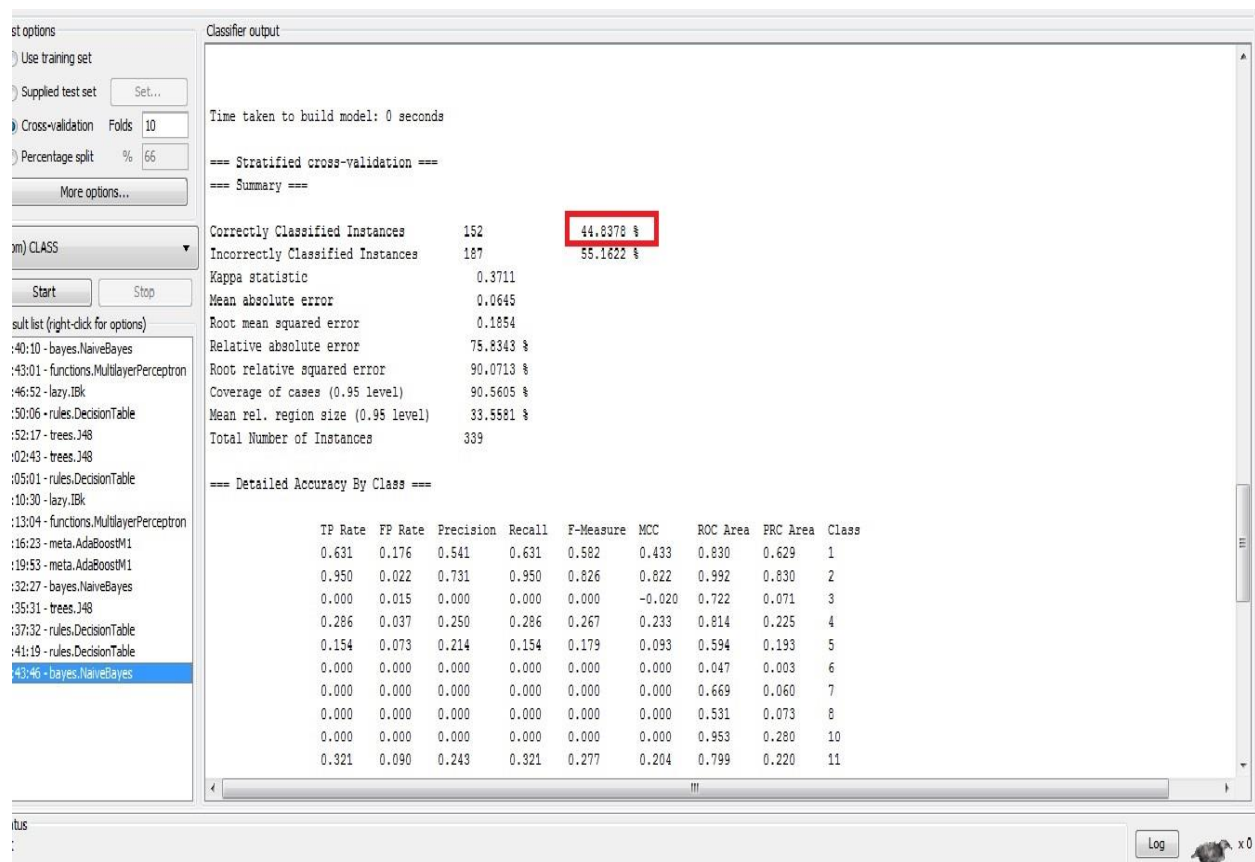
این پروژه را ما به دو صورت پیاده سازی کردیم.

روی دیتاست پیش پردازش انجام دادیم و سپس با انتخاب ویژگی های مختلف و با ابعاد مختلف آن را با الگوریتم *Nave Bayesian* دسته بندی کردیم و میزان دقت آنها را با هم مقایسه کردیم. دیتا ست *Tumor*، دیتاست مقاله می باشد که از نویسنده گرفته شده است. این دیتاست راجع به محل *Tumor* در بدن بیماران می باشد.

روی داده های آموزشی دیتا ست *Tumor*، *feature selection* انجام دادیم، همچنین با الگوریتم *Numerictonominall* پیش پردازش انجام دادیم.



حال با *Nave Bayesian Classifier* این داده ها را کلاس بندی می کنیم و میزان دقت راروی ۳۳۹ نمونه داده و ۱۶ صفت می سنجیم. این کار با ده *fold*، *cross validation* روی کلاس مورد نظر انجام می دهیم. میزان دقت این *classifier* ۴۴٪ است.



اگر بخواهیم قبل از دسته بندی صفات از روش های *feature selection* استفاده کنیم، به این صورت پیش می رود که ابتدا دو الگوریتم جستجو و ارزیاب را تعیین می کنیم.

از الگوریتم *greedy stepwise* یا همان *Wrapper* و *Ranker* یا همان *Filter* استفاده می شود.

روی دیتاست پیش پردازش انجام دادیم و سپس با انتخاب الگوریتم های مختلف آنها را با هم مقایسه کردیم و بهترین الگوریتم بین آنها را یافتیم.

FSS فیلتر:

ابتدا در فیلتر الگوریتم *Ingrain* را به عنوان الگوریتم جستجو انتخاب می کنیم. این عمل با *Cross-Fold 10* *validation* بر روی صفت *Class* انجام می شود. نتیجه به اینصورت می باشد.

Weka Explorer

Preprocess | Classify | Cluster | Associate | Select attributes | Visualize

Attribute Evaluator
Choose **InfoGainAttributeEval**

Search Method
Choose **Ranker -T -1.7976931348623157E308 -N -1**

Attribute Selection Mode
☐ Use full training set
☒ Cross-validation Folds **10**
 Seed **1**

(Nom) CLASS

Start Stop

Result list (right-click for options)

15:36:45 - Ranker + InfoGainAttributeEval
 15:37:03 - Ranker + InfoGainAttributeEval
 15:54:50 - GreedyStepwise + CfsSubsetEv
 17:18:17 - Ranker + InfoGainAttributeEval

Attribute selection output

brain
 skin
 neck
 supraclavicular
 axillar
 mediastinum
 abdominal

Evaluation mode: 10-fold cross-validation

=== Attribute selection 10 fold cross-validation (stratified), seed: 1 ===

average merit	average rank	attribute
0.326 +- 0.009	1 +- 0	3 SEX
0.294 +- 0.008	2 +- 0	12 neck
0.242 +- 0.014	3.3 +- 0.64	14 axillar
0.224 +- 0.014	4.2 +- 0.75	8 peritoneum
0.216 +- 0.008	4.9 +- 0.83	4 bone
0.204 +- 0.013	6 +- 0.63	9 liver
0.189 +- 0.015	6.8 +- 1.08	15 mediastinum
0.175 +- 0.011	8.3 +- 0.64	16 abdominal
0.163 +- 0.01	8.5 +- 0.67	2 AGE
0.13 +- 0.008	10.2 +- 0.4	13 supraclavicular
0.106 +- 0.015	10.8 +- 0.4	6 lunge
0.073 +- 0.008	12.4 +- 0.49	7 pleura
0.069 +- 0.004	12.6 +- 0.49	10 brain
0.056 +- 0.004	14 +- 0	11 skin
0.022 +- 0.002	15 +- 0	5 bone-marrow

و این *FSS* فیلتر را برای دسته بندهای مختلف امتحان می کنیم تا بهترین دسته بند را از بین نتایج انتخاب کنیم.

ابتدا دسته بند *Nave Bayes* را انتخاب می کنیم. نتیجه با همان *Fold 10* به این صورت می باشد. میزان دقت *classifier* ۴۴٪ است.

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier

Choose **NaiveBayes**

Test options

☐ Use training set

☐ Supplied test set Set...

☒ Cross-validation Folds 10

☐ Percentage split % 66

More options...

(Nom) CLASS

Start Stop

Result list (right-click for options)

- 15:40:10 - bayes.NaiveBayes
- 15:43:01 - functions.MultilayerPerceptron
- 15:46:52 - lazy.IBk
- 15:50:06 - rules.DecisionTable
- 15:52:17 - trees.J48
- 17:02:43 - trees.J48
- 17:05:01 - rules.DecisionTable
- 17:10:30 - lazy.IBk
- 17:13:04 - functions.MultilayerPerceptron
- 17:16:23 - meta.AdaBoostM1
- 17:19:53 - meta.AdaBoostM1
- 17:32:27 - bayes.NaiveBayes

Classifier output

1	22.0	1.0	4.0	2.0	19.0	1.0	10.0	2.0	1.0	19.0	13.0	4.0	7.0	2
2	64.0	21.0	7.0	14.0	22.0	2.0	6.0	6.0	3.0	11.0	5.0	5.0	19.0	2
[total]	86.0	22.0	11.0	16.0	41.0	3.0	16.0	8.0	4.0	30.0	18.0	9.0	26.0	4

Time taken to build model: 0 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 152 44.8378 %

Incorrectly Classified Instances 187 55.1622 %

Kappa statistic 0.3711

Mean absolute error 0.0645

Root mean squared error 0.1854

Relative absolute error 75.8343 %

Root relative squared error 90.0713 %

Coverage of cases (0.95 level) 90.5605 %

Mean rel. region size (0.95 level) 33.5581 %

Total Number of Instances 339

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.631	0.176	0.541	0.631	0.582	0.433	0.830	0.629	1
	0.950	0.022	0.731	0.950	0.826	0.822	0.992	0.830	2
	0.000	0.015	0.000	0.000	0.000	-0.020	0.722	0.071	3
	0.286	0.037	0.250	0.286	0.267	0.233	0.814	0.225	4
	0.154	0.073	0.214	0.154	0.179	0.093	0.594	0.193	5

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier

Choose **NaiveBayes**

Test options

☐ Use training set

☐ Supplied test set Set...

☒ Cross-validation Folds 10

☐ Percentage split % 66

More options...

(Nom) CLASS

Start Stop

Result list (right-click for options)

- 15:40:10 - bayes.NaiveBayes

Classifier output

	0.000	0.000	0.000	0.000	0.000	0.000	0.713	0.040	19
	0.000	0.003	0.000	0.000	0.000	-0.004	0.430	0.008	20
	0.000	0.006	0.000	0.000	0.000	-0.004	0.636	0.008	21
	0.782	0.016	0.782	0.782	0.782	0.776	0.868	0.789	22
Weighted Avg.	0.448	0.076	0.373	0.448	0.402	0.340	0.795	0.403	

=== Confusion Matrix ===

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	<-- classified as
53	2	2	6	6	0	0	0	0	4	3	0	2	0	0	1	2	0	0	1	2	1	a = 1
0	19	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	b = 2
5	1	0	1	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	c = 3
4	1	1	4	0	0	0	0	0	0	1	0	3	0	0	0	0	0	0	0	0	0	d = 4
9	0	0	0	6	0	0	0	0	6	4	0	1	0	0	1	8	0	1	0	3	1	e = 5
0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	f = 6
2	0	0	0	3	0	0	0	0	2	5	0	1	0	0	0	1	0	0	0	0	0	g = 7
2	1	0	0	0	0	0	0	0	2	0	0	1	0	0	0	0	0	0	0	0	0	h = 8
1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	i = 10
5	0	0	0	2	0	0	0	0	9	5	0	0	0	0	0	7	0	0	0	0	0	j = 11
1	0	0	0	1	0	0	0	0	2	10	0	0	0	0	0	1	0	0	1	0	1	k = 12
1	0	0	1	2	0	0	0	0	1	0	1	0	0	0	0	1	0	0	0	0	0	l = 13
7	0	1	1	3	0	0	0	0	2	1	0	7	0	0	1	1	0	0	0	0	0	m = 14
1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	n = 15
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	o = 16
3	1	0	0	1	0	0	0	0	2	0	0	3	0	0	0	0	0	0	0	0	0	p = 17
2	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	25	0	0	0	0	0	q = 18
0	0	0	1	2	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	r = 19
0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	s = 20
0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	t = 21

دسته بند AdoboostM1 را انتخاب می کنیم. نتایج بصورت زیر می باشد. میزان دقت classifier ۲۳٪ است.

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier

Choose **AdaBoostM1** -P 100 -S 1 -I 10 -W weka.classifiers.trees.DecisionStump

Test options

☐ Use training set

☐ Supplied test set Set...

☒ Cross-validation Folds 10

☐ Percentage split % 66

More options...

(Nom) CLASS

Start Stop

Result list (right-click for options)

15:40:10 - bayes.NaiveBayes

15:43:01 - functions.MultilayerPerceptron

15:46:52 - lazy.IBk

15:50:06 - rules.DecisionTable

15:52:17 - trees.J48

17:02:43 - trees.J48

17:05:01 - rules.DecisionTable

17:10:30 - lazy.IBk

17:13:04 - functions.MultilayerPerceptron

17:16:23 - meta.AdaBoostM1

17:19:53 - meta.AdaBoostM1

Classifier output

1 2 3 4 5 6 7 8 9 10 11 12 13

0.24778761061946902 0.058997050147492625 0.02654867256637168 0.04129793510324484

Time taken to build model: 0 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 78 23.0088 %

Incorrectly Classified Instances 261 76.9912 %

Kappa statistic 0.0383

Mean absolute error 0.0825

Root mean squared error 0.2033

Relative absolute error 96.8721 %

Root relative squared error 98.7444 %

Coverage of cases (0.95 level) 93.8053 %

Mean rel. region size (0.95 level) 54.6285 %

Total Number of Instances 339

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0.762	0.682	0.269	0.762	0.398	0.075	0.588	0.300	
0.000	0.000	0.000	0.000	0.000	0.000	0.630	0.081	
0.000	0.000	0.000	0.000	0.000	0.000	0.537	0.030	
0.000	0.000	0.000	0.000	0.000	0.000	0.487	0.041	
0.000	0.000	0.000	0.000	0.000	0.000	0.456	0.103	

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier

Choose **AdaBoostM1** -P 100 -S 1 -I 10 -W weka.classifiers.trees.DecisionStump

Test options

☐ Use training set

☐ Supplied test set Set...

☒ Cross-validation Folds 10

☐ Percentage split % 66

More options...

(Nom) CLASS

Start Stop

Result list (right-click for options)

15:40:10 - bayes.NaiveBayes

15:43:01 - functions.MultilayerPerceptron

15:46:52 - lazy.IBk

15:50:06 - rules.DecisionTable

15:52:17 - trees.J48

17:02:43 - trees.J48

17:05:01 - rules.DecisionTable

17:10:30 - lazy.IBk

17:13:04 - functions.MultilayerPerceptron

17:16:23 - meta.AdaBoostM1

17:19:53 - meta.AdaBoostM1

Classifier output

Total number of instances 339

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
0.762	0.682	0.269	0.762	0.398	0.075	0.588	0.300	1	
0.000	0.000	0.000	0.000	0.000	0.000	0.630	0.081	2	
0.000	0.000	0.000	0.000	0.000	0.000	0.537	0.030	3	
0.000	0.000	0.000	0.000	0.000	0.000	0.487	0.041	4	
0.000	0.000	0.000	0.000	0.000	0.000	0.456	0.103	5	
0.000	0.000	0.000	0.000	0.000	0.000	0.265	0.003	6	
0.000	0.000	0.000	0.000	0.000	0.000	0.362	0.032	7	
0.000	0.000	0.000	0.000	0.000	0.000	0.417	0.018	8	
0.000	0.000	0.000	0.000	0.000	0.000	0.576	0.010	10	
0.000	0.000	0.000	0.000	0.000	0.000	0.429	0.070	11	
0.000	0.000	0.000	0.000	0.000	0.000	0.686	0.071	12	
0.000	0.000	0.000	0.000	0.000	0.000	0.202	0.014	13	
0.000	0.000	0.000	0.000	0.000	0.000	0.540	0.079	14	
0.000	0.000	0.000	0.000	0.000	0.000	0.572	0.010	15	
0.000	0.000	0.000	0.000	0.000	0.000	0.286	0.003	16	
0.000	0.000	0.000	0.000	0.000	0.000	0.743	0.055	17	
0.483	0.281	0.139	0.483	0.215	0.124	0.726	0.142	18	
0.000	0.000	0.000	0.000	0.000	0.000	0.633	0.024	19	
0.000	0.000	0.000	0.000	0.000	0.000	0.530	0.009	20	
0.000	0.000	0.000	0.000	0.000	0.000	0.265	0.003	21	
0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	22	
Weighted Avg.	0.230	0.193	0.078	0.230	0.117	0.029	0.558	0.132	

=== Confusion Matrix ===

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	<-- classified as
64	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	a = 1

Status OK

Log

دسته بند Descision Table را انتخاب می کنیم. نتایج بصورت زیر می باشد. میزان دقت classifier 36٪ است.

The screenshot shows the Weka GUI with the Classifier tab active. The classifier selected is DecisionTable. The test options are set to Cross-validation with 10 folds. The classifier output is displayed on the right, showing evaluation metrics for cross-validation. The accuracy is 36.2832%, and the time taken to build the model is 0.16 seconds. The result list on the left shows the classifier was run multiple times, with the last run (17:37:32) selected.

Classifier: DecisionTable -X -S "weka.attributeSelection.BestFirst -D 1 -N 5"

Test options:

- ☐ Use training set
- ☐ Supplied test set
- ☒ Cross-validation Folds: 10
- ☐ Percentage split %: 66

Classifier output:

```

State search after 5 node expansions
Total number of subsets evaluated: 109
Merit of best subset found: 40.413
Evaluation (for feature selection): CV (leave one out)
Feature set: 3,5,8,12,14,15,1

Time taken to build model: 0.16 seconds

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

Correctly Classified Instances      123      36.2832 %
Incorrectly Classified Instances    216      63.7168 %
Kappa statistic                    0.2494
Mean absolute error                 0.083
Root mean squared error             0.1997
Relative absolute error             97.5311 %
Root relative squared error         96.9697 %
Coverage of cases (0.95 level)     94.6903 %
Mean rel. region size (0.95 level) 90.8133 %
Total Number of Instances          339

=== Detailed Accuracy By Class ===

TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC A
0.714    0.353    0.400     0.714   0.513      0.314    0.688
0.750    0.025    0.652     0.750   0.698      0.679    0.965
0.000    0.006    0.000     0.000   0.000      -0.013   0.643
  
```

Result list (right-click for options):

- 15:40:10 - bayes.NaiveBayes
- 15:43:01 - functions.MultilayerPerceptron
- 15:46:52 - lazy.IBk
- 15:50:06 - rules.DecisionTable
- 15:52:17 - trees.J48
- 17:02:43 - trees.J48
- 17:05:01 - rules.DecisionTable
- 17:10:30 - lazy.IBk
- 17:13:04 - functions.MultilayerPerceptron
- 17:16:23 - meta.AdaBoostM1
- 17:19:53 - meta.AdaBoostM1
- 17:32:27 - bayes.NaiveBayes
- 17:35:31 - trees.J48
- 17:37:32 - rules.DecisionTable**

Weka Explorer

PreprocessClassifyClusterAssociateSelect attributesVisualize

Classifier

ChooseDecisionTableC4.5 weka.attributeSelection.BestFirst-D 1-N S

Test options

Use training set

Supplied test setSet...

Cross-validation

Folds10

Percentage split

%66

More options...

(Nom) CLASS

StartStop

Result list (right-click for options)

15:40:10 - bayes.NaiveBayes

15:43:01 - functions.MultilayerPerceptron

15:46:52 - lazy.IBk

15:06:52 - rules.DecisionTable

Classifier output

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
0.714	0.353	0.400	0.714	0.513	0.314	0.688	0.429	1	
0.750	0.025	0.652	0.750	0.698	0.679	0.965	0.544	2	
0.000	0.006	0.000	0.000	0.000	-0.013	0.643	0.066	3	
0.000	0.018	0.000	0.000	0.000	-0.028	0.722	0.114	4	
0.077	0.087	0.103	0.077	0.088	-0.011	0.595	0.136	5	
0.000	0.000	0.000	0.000	0.000	0.000	0.179	0.004	6	
0.071	0.012	0.200	0.071	0.105	0.098	0.625	0.072	7	
0.000	0.000	0.000	0.000	0.000	0.000	0.542	0.025	8	
0.000	0.003	0.000	0.000	0.000	-0.004	0.221	0.005	10	
0.071	0.042	0.133	0.071	0.093	0.040	0.701	0.157	11	
0.438	0.077	0.219	0.438	0.292	0.261	0.816	0.185	12	
0.000	0.000	0.000	0.000	0.000	0.000	0.609	0.038	13	
0.125	0.029	0.250	0.125	0.167	0.134	0.686	0.171	14	
0.000	0.000	0.000	0.000	0.000	0.000	0.325	0.007	15	
0.000	0.000	0.000	0.000	0.000	0.000	0.016	0.003	16	
0.000	0.000	0.000	0.000	0.000	0.000	0.693	0.116	17	
0.690	0.090	0.417	0.690	0.519	0.481	0.856	0.452	18	
0.000	0.000	0.000	0.000	0.000	0.000	0.335	0.016	19	
0.000	0.000	0.000	0.000	0.000	0.000	0.469	0.008	20	
0.000	0.000	0.000	0.000	0.000	0.000	0.964	0.045	21	
Weighted Avg.	0.363	0.118	0.286	0.363	0.303	0.227	0.707	0.282	22

=== Confusion Matrix ===

a b c d e f g h i j k l m n o p q r s t u <-- classified as
60 2 0 2 5 0 1 0 1 1 5 0 2 0 0 0 3 0 0 0 2 a = 1
3 15 2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 b = 2
2 2 0 0 3 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 1 c = 3

StatusOK

Log

3:51 PM2/14/2016

دسته بند IBK را انتخاب می کنیم. نتایج بصورت زیر می باشد. میزان دقت classifier 37٪ است.

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier

Choose **IBK -K 1 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A \"weka.core.EuclideanDistance -R first-last\""**

Test options

☐ Use training set

☐ Supplied test set Set...

☒ Cross-validation Folds 10

☐ Percentage split % 66

More options...

(Nom) CLASS

Start Stop

Result list (right-click for options)

15:40:10 - bayes.NaiveBayes

15:43:01 - functions.MultilayerPerceptron

15:46:52 - lazy.IBK

Classifier output

Test Mode: 10-fold cross-validation

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

IB1 instance-based classifier
using 1 nearest neighbour(s) for classification

Time taken to build model: 0 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	128	37.7581 %
Incorrectly Classified Instances	211	62.2419 %
Kappa statistic	0.2958	
Mean absolute error	0.0649	
Root mean squared error	0.2141	
Relative absolute error	76.1992 %	
Root relative squared error	103.9788 %	
Coverage of cases (0.95 level)	55.7522 %	
Mean rel. region size (0.95 level)	17.8396 %	
Total Number of Instances	339	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.607	0.173	0.537	0.607	0.570	0.418	0.756	0.505	1
	0.900	0.022	0.720	0.900	0.800	0.792	0.984	0.793	2
	0.000	0.030	0.000	0.000	0.000	-0.029	0.564	0.034	3
	0.357	0.025	0.385	0.357	0.370	0.345	0.738	0.156	4
	0.077	0.073	0.120	0.077	0.094	0.004	0.599	0.144	5
	0.000	0.009	0.000	0.000	0.000	-0.005	0.490	0.005	6

Status OK

Log

Classifier

Choose **IBK -K 1 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A \"weka.core.EuclideanDistance -R first-last\""**

Test options

☐ Use training set

☐ Supplied test set Set...

☒ Cross-validation Folds 10

☐ Percentage split % 66

More options...

(Nom) CLASS

Start Stop

Result list (right-click for options)

15:40:10 - bayes.NaiveBayes

15:43:01 - functions.MultilayerPerceptron

15:46:52 - lazy.IBK

Classifier output

Mean rel. region size (0.95 level) 17.8396 %

Total Number of Instances 339

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.607	0.173	0.537	0.607	0.570	0.418	0.756	0.505	1
	0.900	0.022	0.720	0.900	0.800	0.792	0.984	0.793	2
	0.000	0.030	0.000	0.000	0.000	-0.029	0.564	0.034	3
	0.357	0.025	0.385	0.357	0.370	0.345	0.738	0.156	4
	0.077	0.073	0.120	0.077	0.094	0.004	0.599	0.144	5
	0.000	0.009	0.000	0.000	0.000	-0.005	0.490	0.005	6
	0.000	0.037	0.000	0.000	0.000	-0.040	0.549	0.047	7
	0.167	0.009	0.250	0.167	0.200	0.193	0.658	0.104	8
	0.000	0.003	0.000	0.000	0.000	-0.004	0.921	0.135	10
	0.250	0.080	0.219	0.250	0.233	0.160	0.739	0.194	11
	0.438	0.046	0.318	0.438	0.368	0.337	0.767	0.225	12
	0.000	0.009	0.000	0.000	0.000	-0.014	0.579	0.027	13
	0.042	0.060	0.050	0.042	0.045	-0.020	0.657	0.105	14
	0.000	0.003	0.000	0.000	0.000	-0.004	0.856	0.021	15
	0.000	0.003	0.000	0.000	0.000	-0.003	0.618	0.007	16
	0.000	0.030	0.000	0.000	0.000	-0.030	0.762	0.099	17
	0.690	0.058	0.526	0.690	0.597	0.560	0.850	0.428	18
	0.000	0.018	0.000	0.000	0.000	-0.018	0.427	0.016	19
	0.000	0.000	0.000	0.000	0.000	0.000	0.273	0.006	20
	0.000	0.000	0.000	0.000	0.000	0.000	0.169	0.003	21
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	22
Weighted Avg.	0.378	0.076	0.350	0.378	0.360	0.290	0.731	0.323	

دسته بند *Multi layer perceptron* یا *MLP* را انتخاب می کنیم. نتایج بصورت زیر می باشد. میزان دقت *classifier* ۳۸٪ است.

Classifier: Choose **MultilayerPerceptron** 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H a

Test options:
☐ Use training set
☐ Supplied test set Set...
☒ Cross-validation Folds 10
☐ Percentage split % 66
 More options...

(Nom) CLASS
 Start Stop

Result list (right-click for options)
 15:40:10 - bayes.NaiveBayes
 15:43:01 - functions.MultilayerPerceptron

Classifier output
 Class 22
 Input
 Node 20

Time taken to build model: 4.51 seconds

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

==== Summary ====

Correctly Classified Instances	131	38.6431 %
Incorrectly Classified Instances	208	61.3569 %
Kappa statistic	0.3138	
Mean absolute error	0.0622	
Root mean squared error	0.2098	
Relative absolute error	73.0344 %	
Root relative squared error	101.8953 %	
Coverage of cases (0.95 level)	61.3569 %	
Mean rel. region size (0.95 level)	15.8028 %	
Total Number of Instances	339	

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

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
0.583	0.118	0.620	0.583	0.583	0.601	0.476	0.797	0.580	1
0.900	0.025	0.692	0.900	0.900	0.783	0.775	0.978	0.637	2
0.000	0.036	0.000	0.000	0.000	0.000	-0.032	0.598	0.045	3
0.214	0.031	0.231	0.214	0.214	0.222	0.190	0.749	0.208	4
0.179	0.053	0.304	0.179	0.179	0.226	0.160	0.660	0.222	5
0.000	0.009	0.000	0.000	0.000	0.000	-0.005	0.837	0.018	6
0.000	0.046	0.000	0.000	0.000	0.000	-0.045	0.498	0.045	7

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier: Choose **MultilayerPerceptron** 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H a

Test options:
☐ Use training set
☐ Supplied test set Set...
☒ Cross-validation Folds 10
☐ Percentage split % 66
 More options...

(Nom) CLASS
 Start Stop

Result list (right-click for options)
 15:40:10 - bayes.NaiveBayes
 15:43:01 - functions.MultilayerPerceptron

Classifier output
 Mean rel. region size (0.95 level) 15.8028 %
 Total Number of Instances 339

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

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
0.583	0.118	0.620	0.583	0.583	0.601	0.476	0.797	0.580	1
0.900	0.025	0.692	0.900	0.900	0.783	0.775	0.978	0.637	2
0.000	0.036	0.000	0.000	0.000	0.000	-0.032	0.598	0.045	3
0.214	0.031	0.231	0.214	0.214	0.222	0.190	0.749	0.208	4
0.179	0.053	0.304	0.179	0.179	0.226	0.160	0.660	0.222	5
0.000	0.009	0.000	0.000	0.000	0.000	-0.005	0.837	0.018	6
0.000	0.046	0.000	0.000	0.000	0.000	-0.045	0.498	0.045	7
0.000	0.009	0.000	0.000	0.000	0.000	-0.013	0.632	0.071	8
0.000	0.003	0.000	0.000	0.000	0.000	-0.004	0.809	0.030	10
0.214	0.077	0.200	0.214	0.214	0.207	0.133	0.742	0.202	11
0.438	0.059	0.269	0.438	0.438	0.333	0.302	0.821	0.192	12
0.000	0.003	0.000	0.000	0.000	0.000	-0.008	0.486	0.024	13
0.125	0.070	0.120	0.125	0.125	0.122	0.054	0.705	0.173	14
0.000	0.006	0.000	0.000	0.000	0.000	-0.006	0.269	0.006	15
0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.237	0.004	16
0.200	0.046	0.118	0.200	0.200	0.148	0.120	0.724	0.092	17
0.690	0.045	0.588	0.690	0.690	0.635	0.600	0.949	0.510	18
0.000	0.018	0.000	0.000	0.000	0.000	-0.018	0.752	0.040	19
0.000	0.003	0.000	0.000	0.000	0.000	-0.004	0.530	0.010	20
0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.565	0.007	21
0.667	0.019	0.727	0.667	0.667	0.696	0.674	0.808	0.720	22
Weighted Avg.	0.386	0.062	0.382	0.386	0.380	0.318	0.764	0.357	

==== Confusion Matrix ====

a b c d e f g h i j k l m n o p q r s t u <-- classified as

دسته بند J48 را انتخاب می کنیم. نتایج بصورت زیر می باشد. میزان دقت classifier 40٪ است.

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier

Choose **J48 -C 0.25 -M 2**

Test options

☐ Use training set

☐ Supplied test set Set...

☒ Cross-validation Folds **10**

☐ Percentage split % **66**

More options...

(Nom) CLASS

Start Stop

Result list (right-click for options)

15:40:10 - bayes.NaiveBayes

15:43:01 - functions.MultilayerPerceptron

15:46:52 - lazy.IBk

15:50:06 - rules.DecisionTable

15:52:17 - trees.J48

17:02:43 - trees.J48

17:05:01 - rules.DecisionTable

17:10:30 - lazy.IBk

17:13:04 - functions.MultilayerPerceptron

17:16:23 - meta.AdaBoostM1

17:19:53 - meta.AdaBoostM1

17:32:27 - bayes.NaiveBayes

17:35:31 - trees.J48

Classifier output

Size of the tree : 97

Time taken to build model: 0.02 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	138	40.708 %
Incorrectly Classified Instances	201	59.292 %
Kappa statistic	0.3256	
Mean absolute error	0.0645	
Root mean squared error	0.2002	
Relative absolute error	75.7744 %	
Root relative squared error	97.2449 %	
Coverage of cases (0.95 level)	64.8968 %	
Mean rel. region size (0.95 level)	19.7359 %	
Total Number of Instances	339	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.595	0.188	0.510	0.595	0.549	0.388	0.741	0.511	1
	0.950	0.025	0.704	0.950	0.809	0.805	0.959	0.621	2
	0.000	0.015	0.000	0.000	0.000	-0.020	0.484	0.026	3
	0.286	0.025	0.333	0.286	0.308	0.281	0.772	0.197	4
	0.103	0.083	0.138	0.103	0.118	0.022	0.591	0.145	5
	0.000	0.000	0.000	0.000	0.000	0.000	0.473	0.003	6

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier

Choose **J48 -C 0.25 -M 2**

Test options

☐ Use training set

☐ Supplied test set Set...

☒ Cross-validation Folds **10**

☐ Percentage split % **66**

More options...

(Nom) CLASS

Start Stop

Result list (right-click for options)

15:40:10 - bayes.NaiveBayes

15:43:01 - functions.MultilayerPerceptron

15:46:52 - lazy.IBk

15:50:06 - rules.DecisionTable

15:52:17 - trees.J48

Classifier output

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.595	0.188	0.510	0.595	0.549	0.388	0.741	0.511	1
	0.950	0.025	0.704	0.950	0.809	0.805	0.959	0.621	2
	0.000	0.015	0.000	0.000	0.000	-0.020	0.484	0.026	3
	0.286	0.025	0.333	0.286	0.308	0.281	0.772	0.197	4
	0.103	0.083	0.138	0.103	0.118	0.022	0.591	0.145	5
	0.000	0.000	0.000	0.000	0.000	0.000	0.473	0.003	6
	0.000	0.034	0.000	0.000	0.000	-0.038	0.671	0.069	7
	0.000	0.009	0.000	0.000	0.000	-0.013	0.492	0.048	8
	0.000	0.003	0.000	0.000	0.000	-0.004	0.746	0.253	10
	0.179	0.068	0.192	0.179	0.185	0.115	0.693	0.176	11
	0.563	0.065	0.300	0.563	0.391	0.371	0.796	0.179	12
	0.000	0.009	0.000	0.000	0.000	-0.014	0.373	0.021	13
	0.250	0.044	0.300	0.250	0.273	0.224	0.603	0.191	14
	0.000	0.000	0.000	0.000	0.000	0.000	0.476	0.006	15
	0.000	0.000	0.000	0.000	0.000	0.000	0.485	0.003	16
	0.000	0.021	0.000	0.000	0.000	-0.025	0.667	0.072	17
	0.724	0.071	0.488	0.724	0.583	0.549	0.835	0.505	18
	0.000	0.000	0.000	0.000	0.000	0.000	0.429	0.018	19
	0.000	0.000	0.000	0.000	0.000	0.000	0.475	0.006	20
	0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	21
	0.833	0.013	0.833	0.833	0.833	0.831	0.936	0.757	22
Weighted Avg.	0.407	0.080	0.350	0.407	0.372	0.302	0.713	0.330	

=== Confusion Matrix ===

FSS Wrapper ابتدا با *GreedyStepwise* و *CFs subset Eval* کار میکنیم، این عمل با *Cross-Fold* validation بر روی صفت *Class* انجام می شود. نتیجه به اینصورت می باشد.

The screenshot shows the Weka Explorer interface with the following settings:

- Attribute Evaluator:** CfsSubsetEval -P 1 -E 1
- Search Method:** GreedyStepwise -T -1.7 76931348623157E308 -N -1 -num-slots 1
- Attribute Selection Mode:** Cross-validation (selected), Folds: 10, Seed: 1
- Class:** (Nom) CLASS

The **Attribute selection output** pane displays the following results:

```

skin
neck
supraclavicular
axillar
mediastinum
abdominal
Evaluation mode: 10-fold cross-validation

```

The **Result list** shows the execution of GreedyStepwise + CfsSubsetEval. The **Attribute selection 10 fold cross-validation (stratified), seed: 1** results are as follows:

number of folds (%)	attribute
10(100 %)	2 AGE
10(100 %)	3 SEX
9(90 %)	4 bone
0(0 %)	5 bone-marrow
9(90 %)	6 lunge
0(0 %)	7 pleura
10(100 %)	8 peritoneum
10(100 %)	9 liver
0(0 %)	10 brain
1(10 %)	11 skin
10(100 %)	12 neck
3(30 %)	13 supraclavicular
10(100 %)	14 axillar
10(100 %)	15 mediastinum
9(90 %)	16 abdominal

و این *FSS Wrapper* برای دسته بندهای مختلف امتحان می کنیم تا بهترین دسته بند را از بین نتایج انتخاب کنیم.

ابتدا دسته بند *Naive Bayes* را انتخاب می کنیم. نتیجه با همان *Fold 10* به این صورت می باشد. میزان دقت *classifier* ۴۰٪ است.

Classifier output

Time taken to build model: 0.09 seconds

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

Correctly Classified Instances	138	40.708 %
Incorrectly Classified Instances	201	59.292 %
Kappa statistic	0.3256	
Mean absolute error	0.0645	
Root mean squared error	0.2002	
Relative absolute error	75.7744 %	
Root relative squared error	97.2449 %	
Coverage of cases (0.95 level)	64.8968 %	
Mean rel. region size (0.95 level)	19.7359 %	
Total Number of Instances	339	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.595	0.188	0.510	0.595	0.549	0.388	0.741	0.511	1
	0.950	0.025	0.704	0.950	0.809	0.805	0.959	0.621	2
	0.000	0.015	0.000	0.000	0.000	-0.020	0.484	0.026	3
	0.286	0.025	0.333	0.286	0.308	0.281	0.772	0.197	4
	0.103	0.083	0.138	0.103	0.118	0.022	0.591	0.145	5
	0.000	0.000	0.000	0.000	0.000	0.000	0.473	0.003	6
	0.000	0.034	0.000	0.000	0.000	-0.038	0.671	0.069	7

Weka Explorer

Preprocess | **Classify** | Cluster | Associate | Select attributes | Visualize

Classifier: Choose **NaiveBayes**

Test options:
☐ Use training set
☐ Supplied test set (Set...)
☒ Cross-validation Folds: 10
☐ Percentage split %: 66
 More options...

(Nom) CLASS

Start Stop

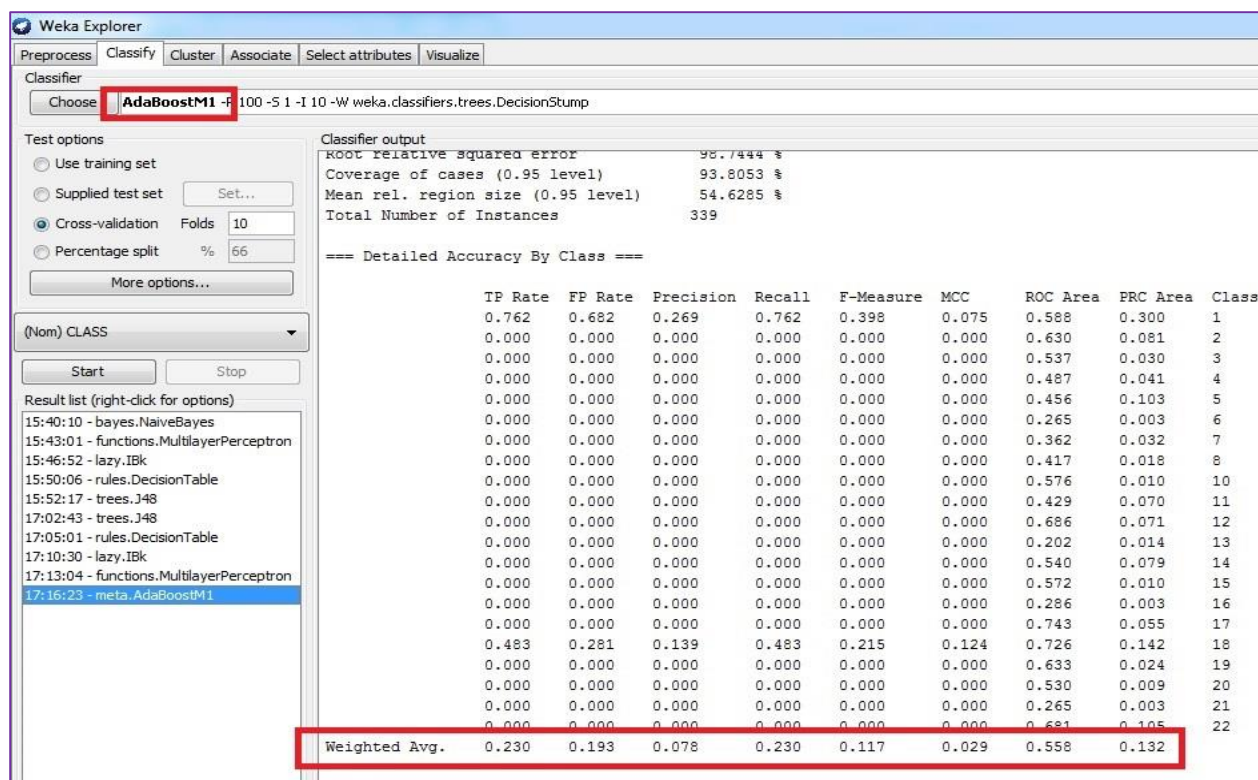
Result list (right-click for options):
 15:40:10 - bayes.NaiveBayes
 15:43:01 - functions.MultilayerPerceptron
 15:46:52 - lazy.IBk
 15:50:06 - rules.DecisionTable
15:52:17 - trees.J48

Classifier output:

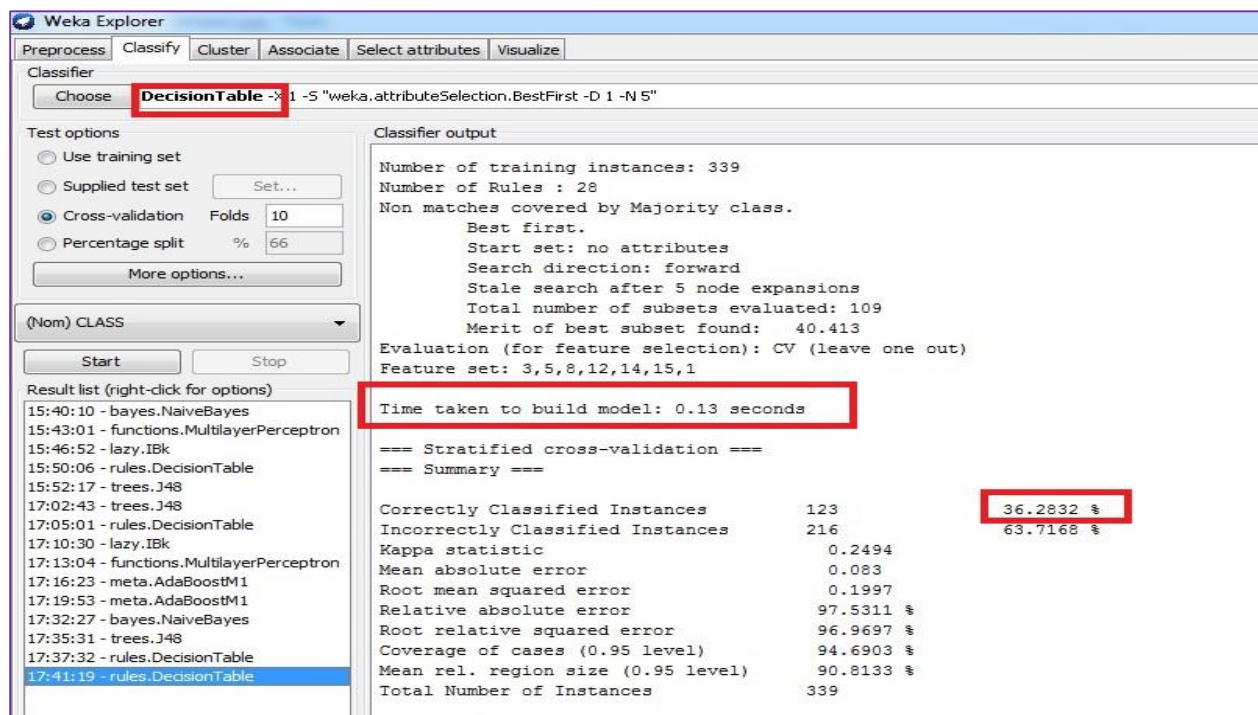
Mean rel. region size (0.95 level): 15.7395 %
 Total Number of Instances: 339

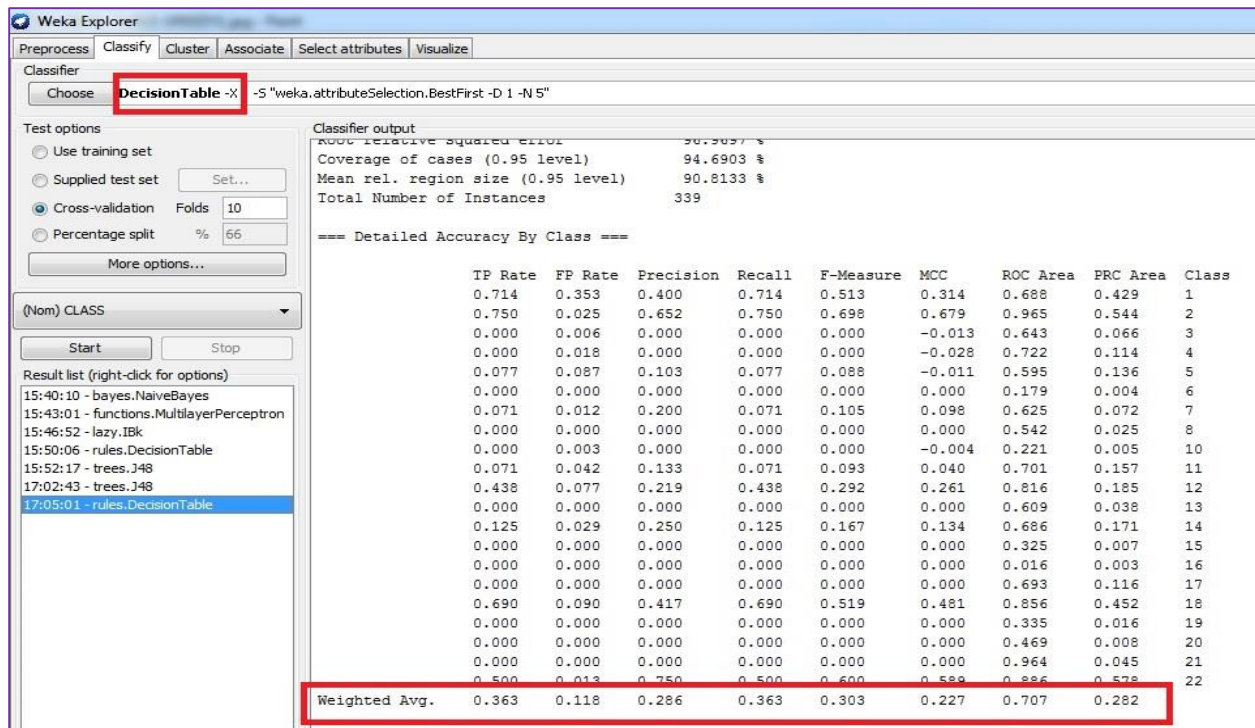
=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
0.595	0.188	0.510	0.595	0.549	0.388	0.741	0.511	1	
0.950	0.025	0.704	0.950	0.809	0.805	0.959	0.621	2	
0.000	0.015	0.000	0.000	0.000	-0.020	0.484	0.026	3	
0.286	0.025	0.333	0.286	0.308	0.281	0.772	0.197	4	
0.103	0.083	0.138	0.103	0.118	0.022	0.591	0.145	5	
0.000	0.000	0.000	0.000	0.000	0.000	0.473	0.003	6	
0.000	0.034	0.000	0.000	0.000	-0.038	0.671	0.069	7	
0.000	0.009	0.000	0.000	0.000	-0.013	0.492	0.048	8	
0.000	0.003	0.000	0.000	0.000	-0.004	0.746	0.253	10	
0.179	0.068	0.192	0.179	0.185	0.115	0.693	0.176	11	
0.563	0.065	0.300	0.563	0.391	0.371	0.796	0.179	12	
0.000	0.009	0.000	0.000	0.000	-0.014	0.373	0.021	13	
0.250	0.044	0.300	0.250	0.273	0.224	0.603	0.191	14	
0.000	0.000	0.000	0.000	0.000	0.000	0.476	0.006	15	
0.000	0.000	0.000	0.000	0.000	0.000	0.485	0.003	16	
0.000	0.021	0.000	0.000	0.000	-0.025	0.667	0.072	17	
0.724	0.071	0.488	0.724	0.583	0.549	0.835	0.505	18	
0.000	0.000	0.000	0.000	0.000	0.000	0.429	0.018	19	
0.000	0.000	0.000	0.000	0.000	0.000	0.475	0.006	20	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	21	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	22	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	23	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	24	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	25	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	26	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	27	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	28	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	29	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	30	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	31	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	32	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	33	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	34	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	35	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	36	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	37	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	38	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	39	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	40	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	41	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	42	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	43	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	44	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	45	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	46	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	47	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	48	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	49	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	50	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	51	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	52	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	53	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	54	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	55	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	56	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	57	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	58	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	59	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	60	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	61	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	62	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	63	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	64	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	65	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	66	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	67	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	68	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	69	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	70	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	71	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	72	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	73	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	74	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	75	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	76	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	77	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	78	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	79	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	80	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	81	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	82	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	83	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	84	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	85	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	86	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	87	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	88	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	89	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	90	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	91	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	92	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	93	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	94	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	95	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	96	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	97	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	98	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	99	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	100	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	101	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	102	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	103	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	104	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	105	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	106	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	107	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	108	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	109	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	110	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	111	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	112	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	113	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	114	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	115	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	116	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	117	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	118	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	119	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	120	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	121	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	122	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	123	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	124	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	125	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	126	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	127	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	128	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	129	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	130	
0.000	0.000	0.000	0.000						

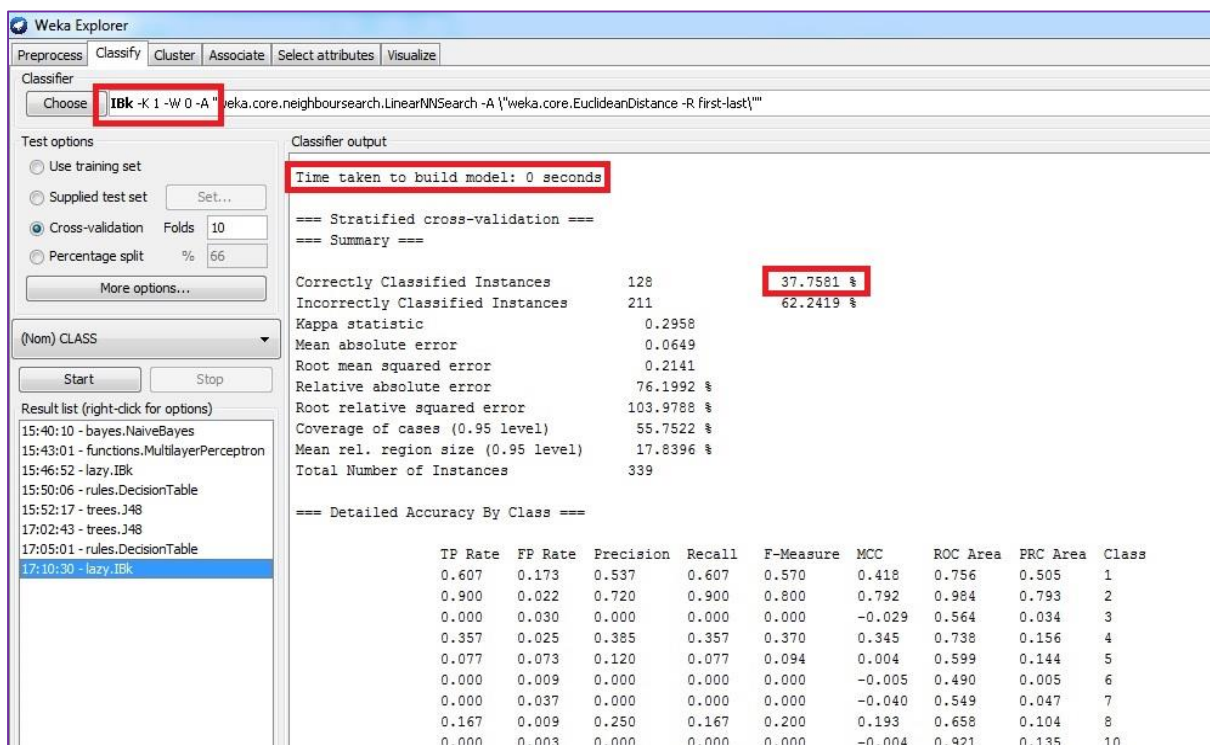


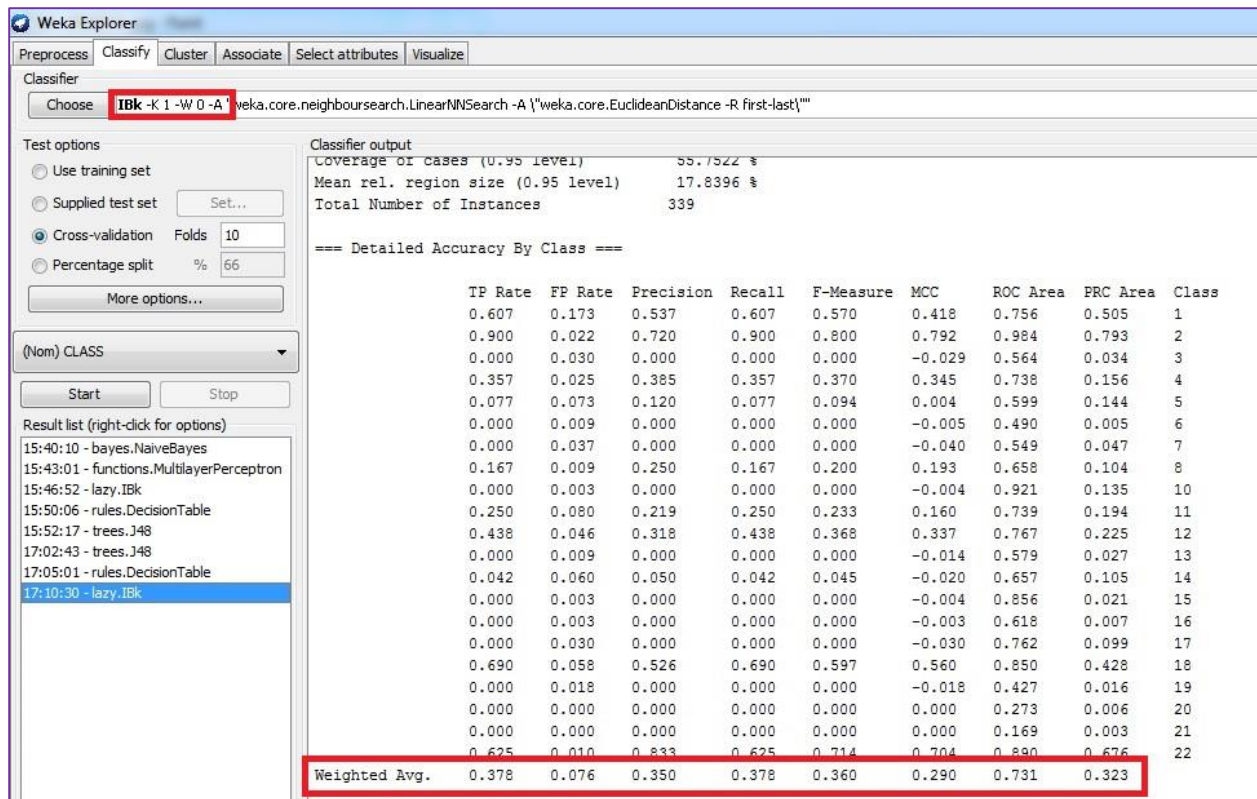
دسته بند *Descision Table* را انتخاب می کنیم. نتایج بصورت زیر می باشد. میزان دقت *classifier* 36٪ است.



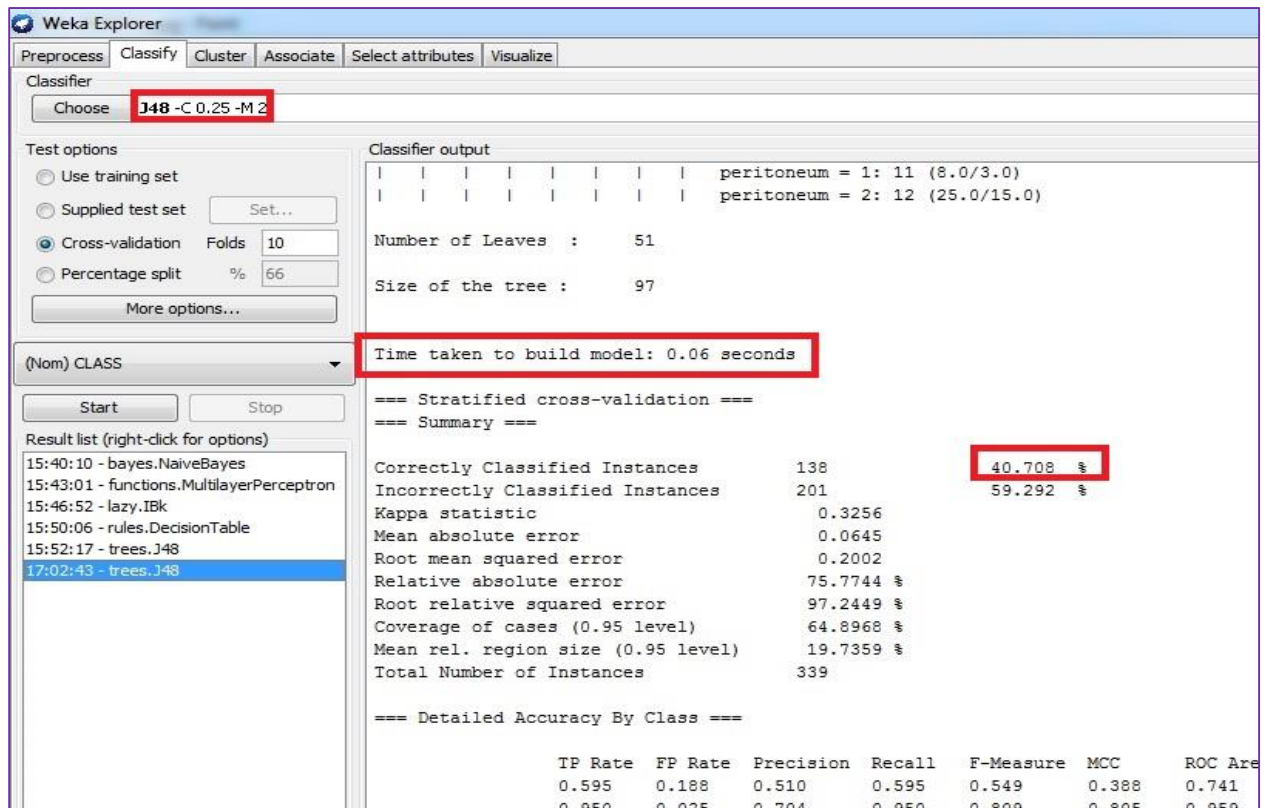


دسته بند IBK را انتخاب می کنیم. نتایج بصورت زیر می باشد. میزان دقت classifier, 37٪ است.





دسته بند 48 را انتخاب می کنیم. نتایج بصورت زیر می باشد. میزان دقت classifier 40٪ است.



Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier: Choose **J48 -C 0.25 -M 2**

Test options:

- Use training set
- Supplied test set: Set...
- Cross-validation** Folds: 10
- Percentage split: % 66

 More options...

(Nom) CLASS

Start Stop

Result list (right-click for options):

- 15:40:10 - bayes.NaiveBayes
- 15:43:01 - functions.MultilayerPerceptron
- 15:46:52 - lazy.IBk
- 15:50:06 - rules.DecisionTable
- 15:52:17 - trees.J48
- 17:02:43 - trees.J48**

Classifier output:

Mean rel. region size (0.95 level) 19.7359 %
Total Number of Instances 339

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
0.595	0.188	0.510	0.595	0.549	0.388	0.741	0.511	1	
0.950	0.025	0.704	0.950	0.809	0.805	0.959	0.621	2	
0.000	0.015	0.000	0.000	0.000	-0.020	0.484	0.026	3	
0.286	0.025	0.333	0.286	0.308	0.281	0.772	0.197	4	
0.103	0.083	0.138	0.103	0.118	0.022	0.591	0.145	5	
0.000	0.000	0.000	0.000	0.000	0.000	0.473	0.003	6	
0.000	0.034	0.000	0.000	0.000	-0.038	0.671	0.069	7	
0.000	0.009	0.000	0.000	0.000	-0.013	0.492	0.048	8	
0.000	0.003	0.000	0.000	0.000	-0.004	0.746	0.253	10	
0.179	0.068	0.192	0.179	0.185	0.115	0.693	0.176	11	
0.563	0.065	0.300	0.563	0.391	0.371	0.796	0.179	12	
0.000	0.009	0.000	0.000	0.000	-0.014	0.373	0.021	13	
0.250	0.044	0.300	0.250	0.273	0.224	0.603	0.191	14	
0.000	0.000	0.000	0.000	0.000	0.000	0.476	0.006	15	
0.000	0.000	0.000	0.000	0.000	0.000	0.485	0.003	16	
0.000	0.021	0.000	0.000	0.000	-0.025	0.667	0.072	17	
0.724	0.071	0.488	0.724	0.583	0.549	0.835	0.505	18	
0.000	0.000	0.000	0.000	0.000	0.000	0.429	0.018	19	
0.000	0.000	0.000	0.000	0.000	0.000	0.475	0.006	20	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	21	
0.000	0.000	0.000	0.000	0.000	0.000	0.494	0.003	22	
Weighted Avg.	0.407	0.080	0.350	0.407	0.372	0.302	0.713	0.330	

دسته بند *Multi layer perceptron* یا *MLP* را انتخاب می کنیم. نتایج بصورت زیر می باشد. میزان دقت *classifier* ۳۸٪ است.

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier: Choose **MultilayerPerceptron -L .3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H a**

Test options:

- Use training set
- Supplied test set: Set...
- Cross-validation** Folds: 10
- Percentage split: % 66

 More options...

(Nom) CLASS

Start Stop

Result list (right-click for options):

- 15:40:10 - bayes.NaiveBayes
- 15:43:01 - functions.MultilayerPerceptron
- 15:46:52 - lazy.IBk
- 15:50:06 - rules.DecisionTable
- 15:52:17 - trees.J48
- 17:02:43 - trees.J48
- 17:05:01 - rules.DecisionTable
- 17:10:30 - lazy.IBk
- 17:13:04 - functions.MultilayerPerceptron**

Classifier output:

node 20

Time taken to build model: 4.3 seconds

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

Correctly Classified Instances 131 **38.6431 %**
Incorrectly Classified Instances 208 61.3569 %

Kappa statistic 0.3138
Mean absolute error 0.0622
Root mean squared error 0.2098
Relative absolute error 73.0344 %
Root relative squared error 101.8953 %
Coverage of cases (0.95 level) 61.3569 %
Mean rel. region size (0.95 level) 15.8028 %
Total Number of Instances 339

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
0.583	0.118	0.620	0.583	0.601	0.476	0.797	0.580	1	
0.900	0.025	0.692	0.900	0.783	0.775	0.978	0.637	2	
0.000	0.036	0.000	0.000	0.000	-0.032	0.598	0.045	3	
0.214	0.031	0.231	0.214	0.222	0.190	0.749	0.208	4	
0.179	0.053	0.304	0.179	0.226	0.160	0.660	0.222	5	
0.000	0.009	0.000	0.000	0.000	-0.005	0.837	0.018	6	

Weka Explorer

Preprocess | Classify | Cluster | Associate | Select attributes | Visualize

Classifier: Choose **MultilayerPerceptron** -L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H a

Test options:
☐ Use training set
☐ Supplied test set (Set...)
☒ Cross-validation Folds: 10
☐ Percentage split %: 66
 More options...

(Nom) CLASS

Start Stop

Result list (right-click for options):
 15:40:10 - bayes.NaiveBayes
 15:43:01 - functions.MultilayerPerceptron
 15:46:52 - lazy.IBk
 15:50:06 - rules.DecisionTable
 15:52:17 - trees.J48
 17:02:43 - trees.J48
 17:05:01 - rules.DecisionTable
 17:10:30 - lazy.IBk
 17:13:04 - functions.MultilayerPerceptron

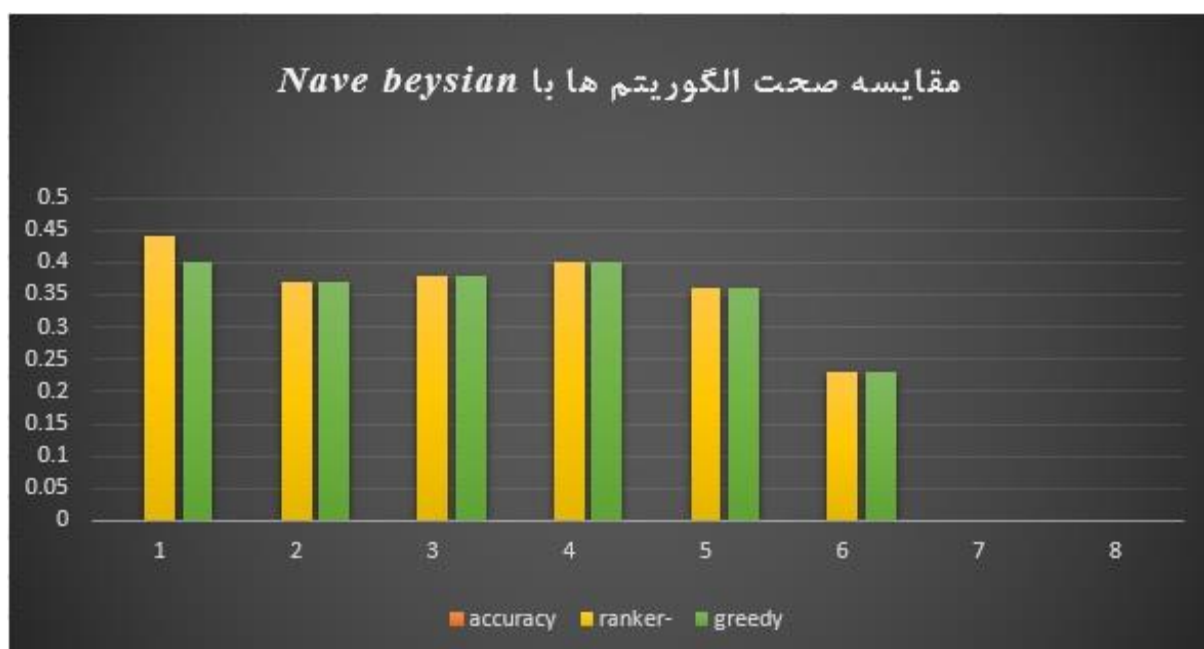
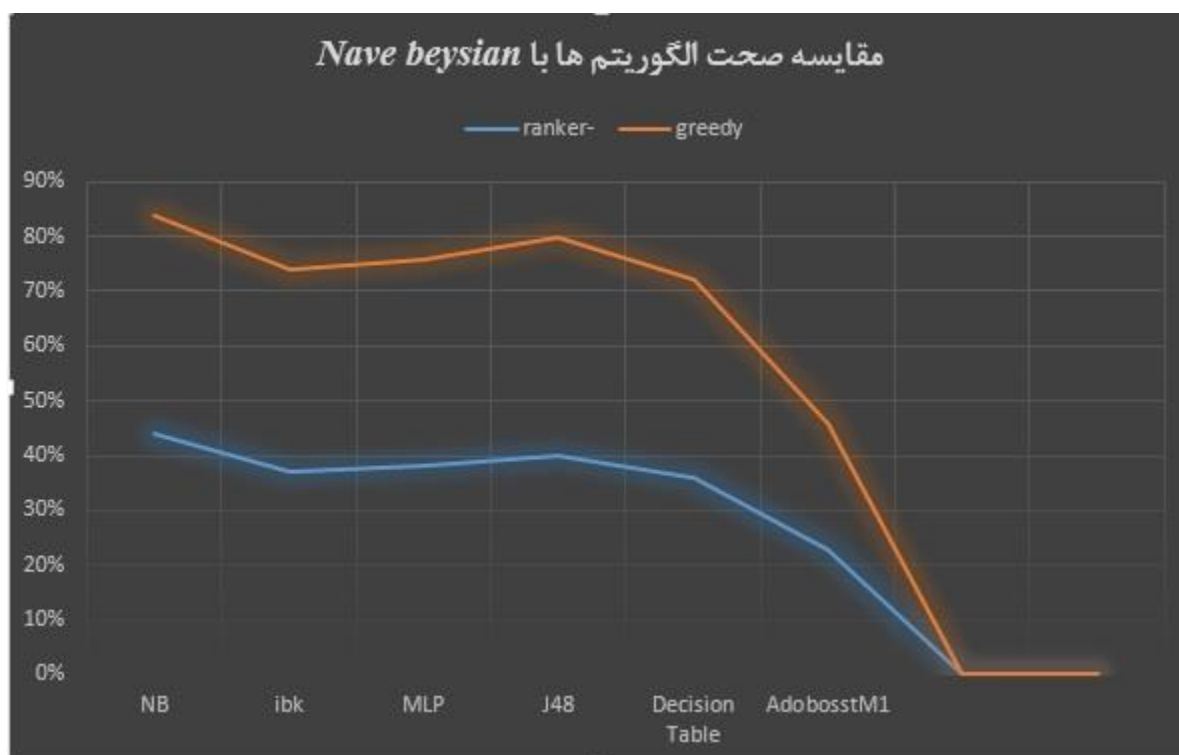
Classifier output:
 Root relative squared error: 101.8953 %
 Coverage of cases (0.95 level): 61.3569 %
 Mean rel. region size (0.95 level): 15.8028 %
 Total Number of Instances: 339

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
0.583	0.118	0.620	0.583	0.601	0.476	0.797	0.580	1	
0.900	0.025	0.692	0.900	0.783	0.775	0.978	0.637	2	
0.000	0.036	0.000	0.000	0.000	-0.032	0.598	0.045	3	
0.214	0.031	0.231	0.214	0.222	0.190	0.749	0.208	4	
0.179	0.053	0.304	0.179	0.226	0.160	0.660	0.222	5	
0.000	0.009	0.000	0.000	0.000	-0.005	0.837	0.018	6	
0.000	0.046	0.000	0.000	0.000	-0.045	0.498	0.045	7	
0.000	0.009	0.000	0.000	0.000	-0.013	0.632	0.071	8	
0.000	0.003	0.000	0.000	0.000	-0.004	0.809	0.030	10	
0.214	0.077	0.200	0.214	0.207	0.133	0.742	0.202	11	
0.438	0.059	0.269	0.438	0.333	0.302	0.821	0.192	12	
0.000	0.003	0.000	0.000	0.000	-0.008	0.486	0.024	13	
0.125	0.070	0.120	0.125	0.122	0.054	0.705	0.173	14	
0.000	0.006	0.000	0.000	0.000	-0.006	0.269	0.006	15	
0.000	0.000	0.000	0.000	0.000	0.000	0.237	0.004	16	
0.200	0.046	0.118	0.200	0.148	0.120	0.724	0.092	17	
0.690	0.045	0.588	0.690	0.635	0.600	0.949	0.510	18	
0.000	0.018	0.000	0.000	0.000	-0.018	0.752	0.040	19	
0.000	0.003	0.000	0.000	0.000	-0.004	0.530	0.010	20	
0.000	0.000	0.000	0.000	0.000	0.000	0.565	0.007	21	
0.662	0.030	0.708	0.662	0.606	0.654	0.900	0.530	22	
Weighted Avg.	0.386	0.062	0.382	0.386	0.380	0.318	0.764	0.357	

در این قسمت مقدار صحت عملکرد این classifier ها را باهم مورد مقایسه قرار می دهیم.

accuracy	NB	ibk	MLP	J48	Decision Table	AdobosstM1
ranker-	44%	37%	38%	40%	36%	23%
greedy	40%	37%	38%	40%	36%	23%



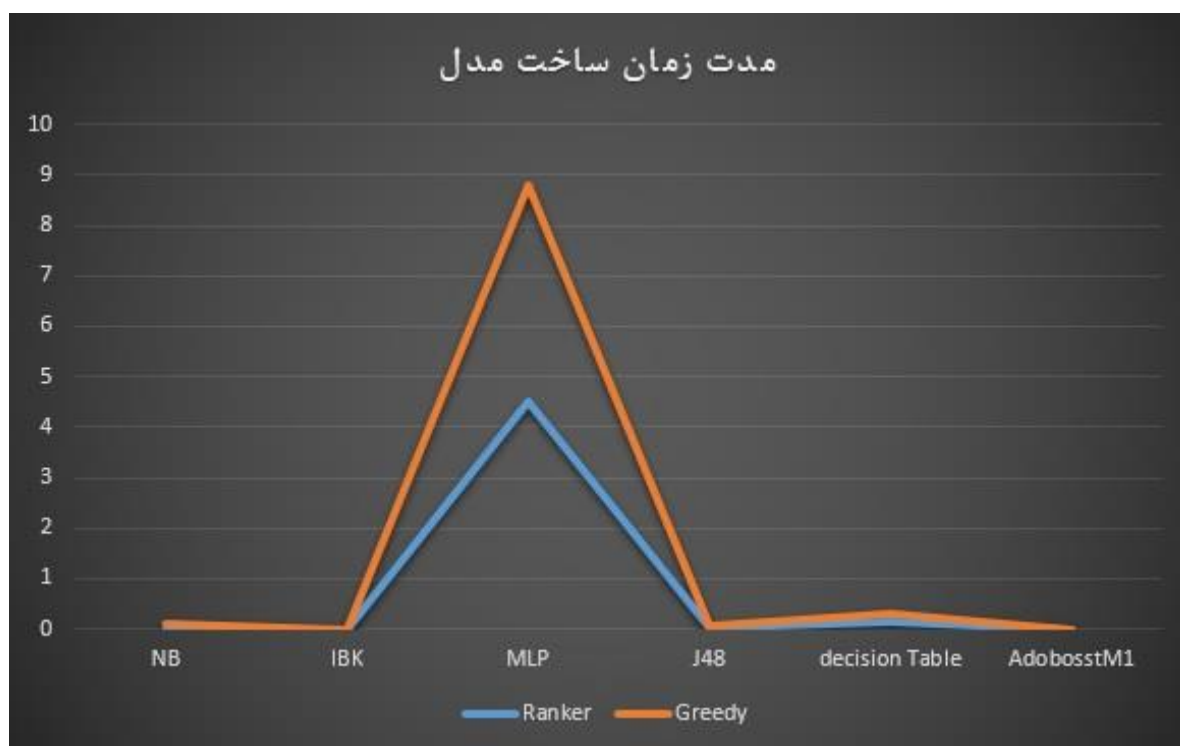
برای بدست آوردن نرخ خطا برای هر دسته بند از فرمول زیر استفاده می کنیم که این نرخ خطا برای *Nave* *Beysian* مناسب نیست اما نسبت به دیگر دسته بند ها نرخ خطای کمتر دارد و همچنین زمان کمتری برای محاسبه صرف می کند.

Error Rate=100- Accuracy

Error Rate-Nave Beysian=100- 44= 56%

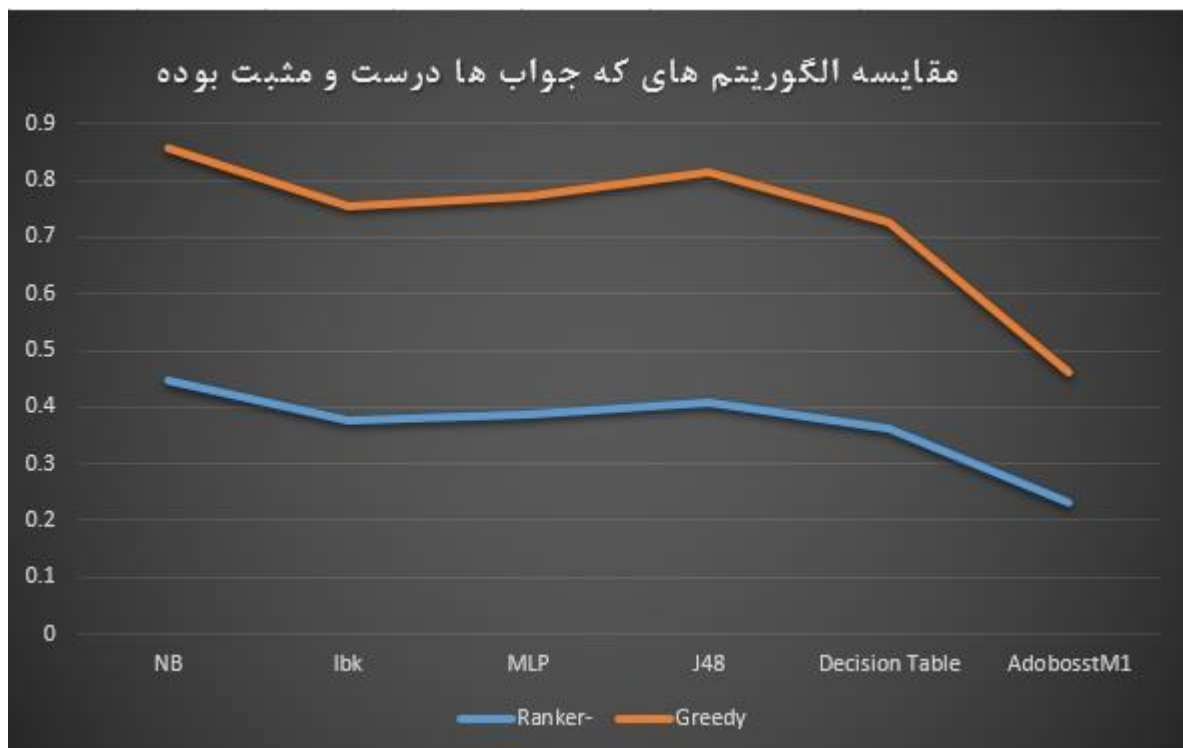
بنابراین صحت عملکرد *Nave Bayesian* از همه بالاتر است. بنابراین نتایج آن برای بررسی افراد با بیماری *Tumor* قابل اعتمادتر می باشد. همان گونه که مشاهده می شود فیلتر نسبت به *Wrapper* بهتر جواب می دهد. مدت زمانی که هر دسته بند طول می کشد تا مدل را بسازد به صورت زیر می باشد. همان طور که مشاهده می کنید فیلتر *Nave Beysian* هیچ مدت زمانی را صرف نمی کند، همچنین مشاهده کردیم دسته بند J48 صحت عملکرد مناسبی دارد، اما زمان ساخت مدل آن بیشتر است.

Time	NB	IBK	MLP	J48	decision Table	AdobosstM1
Ranker	0	0	4.51	0.02	0.16	0
Greedy	0.09	0	4.3	0.06	0.13	0



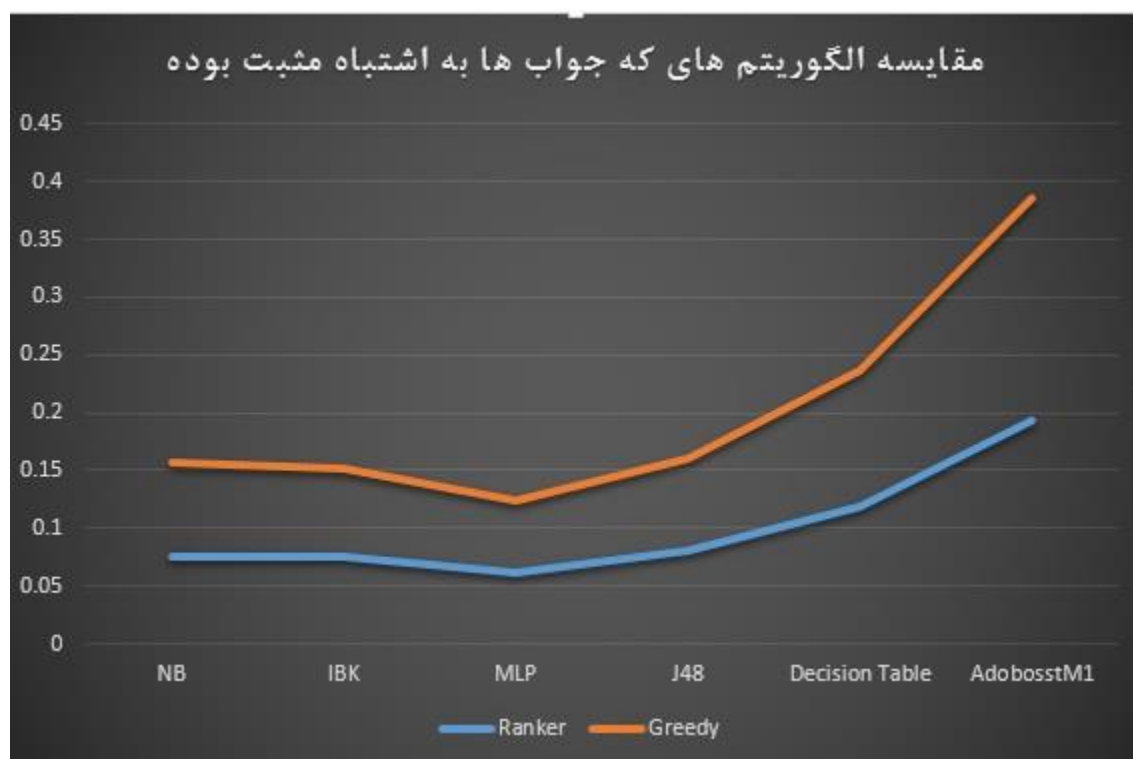
در این قسمت افرادی که بیمار هستن و جواب آزمایش آن ها نیز مثبت بوده (یعنی درست تشخیص داده شده است) را مشاهده می کنید. *Nave Bayesian* بیشترین تشخیص را می دهد.

TP	NB	lbk	MLP	J48	Decision Table	AdobosstM1
Ranker-	0.448	0.378	0.386	0.407	0.363	0.23
Greedy	0.407	0.378	0.386	0.407	0.363	0.23



در این قسمت افرادی که بیمار نبودند ولی جواب آزمایش آن ها مثبت بوده است (یعنی اشتباه تشخیص داده است). در اینجا شبکه عصبی کمترین اشتباه را داشته است.

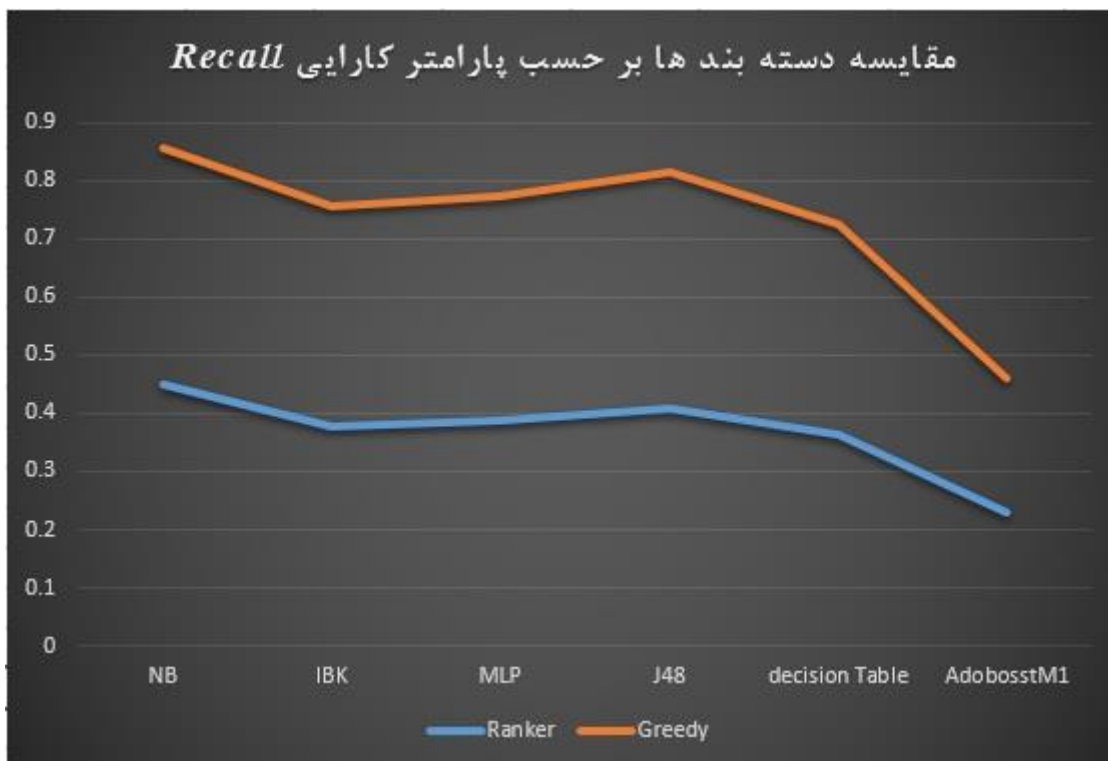
FP	NB	IBK	MLP	J48	Decision Table	AdobosstM1
Ranker	0.076	0.076	0.062	0.08	0.118	0.193
Greedy	0.08	0.076	0.062	0.08	0.118	0.193



در این قسمت پارامتر کارایی Recall را مشاهده می کنید که این کارایی از فرمول زیر بدست می آید.

$$Sensitivity = \frac{TP}{TP + FN} = TP Rate$$

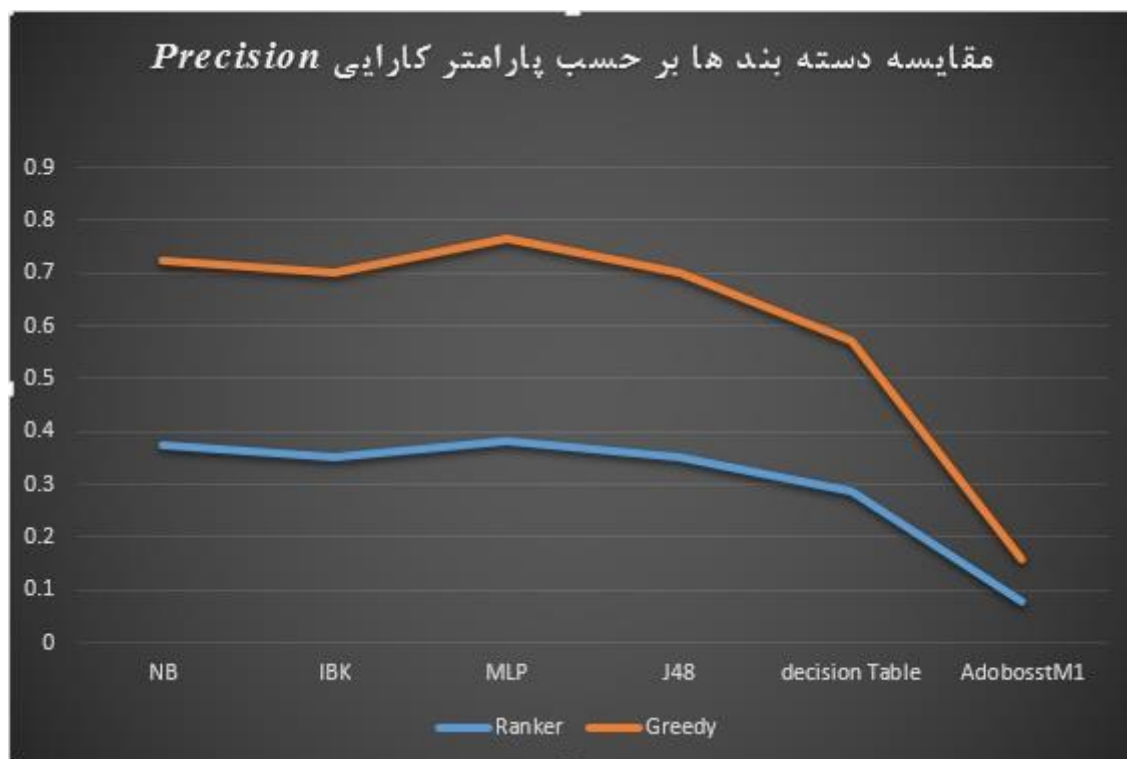
Recall	NB	IBK	MLP	J48	decision Table	AdobosstM1
Ranker	0.448	0.378	0.386	0.407	0.363	0.23
Greedy	0.407	0.378	0.386	0.407	0.363	0.23



در این قسمت پارامتر کارایی Precision را مشاهده می کنید که این کارایی از فرمول زیر بدست می آید.

$$Specificity = \frac{TN}{TN + FP} = 1 - FP Rate$$

Precision	NB	IBK	MLP	J48	decision Table	AdobosstM1
Ranker	0.373	0.35	0.382	0.35	0.286	0.078
Greedy	0.35	0.35	0.382	0.35	0.286	0.078



در این قسمت پارامتر کارایی *Precision* را مشاهده می کنید.

Roc Area	NB	IBK	MLP	J48	decision Table	AdobosstM1
Ranker	0.795	0.731	0.764	0.713	0.707	0.558
Greedy	0.713	0.731	0.764	0.713	0.707	0.558

