Name: Vinayak Soni
Roll No: 67

Practical No. 5

1. Press the Explorer button on the main panel and load the weather dataset and answer the following questions

(a) How many instances are there in the dataset?

Current relation

Relation: weather.symbolic
Instances: 14
Attributes: 5
Sum of weights: 14

(b) State the names of the attributes along with their types and values.

Selected attribute

Name: temperature
Missing: 0 (0%)
Distinct: 3
Type: Nominal
Unique: 0 (0%)

No.	Label	Count	Weight
1	hot	4	4
2	mild	6	6
3	cool	4	4

Selected attribute

Name: outlook
Missing: 0 (0%)
Distinct: 3
Type: Nominal
Unique: 0 (0%)

No.	Label	Count	Weight
1	sunny	5	5
2	overcast	4	4
3	rainy	5	5

Selected attribute

Name: temperature
Missing: 0 (0%)
Distinct: 3
Type: Nominal
Unique: 0 (0%)

No.	Label	Count	Weight
1	hot	4	4
2	mild	6	6
3	cool	4	4

Selected attribute

Name: windy Type: Nominal
Missing: 0 (0%) Distinct: 2 Unique: 0 (0%)

No.	Label	Count	Weight
1	TRUE	6	6
2	FALSE	8	8

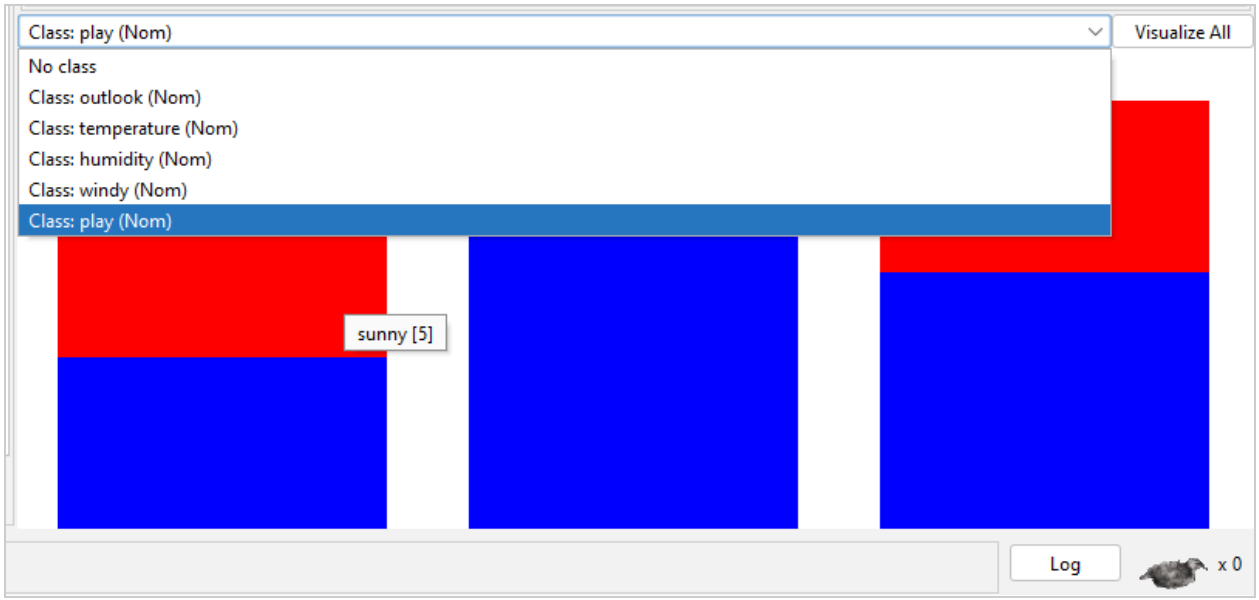
(c) What is the class attribute?

Selected attribute

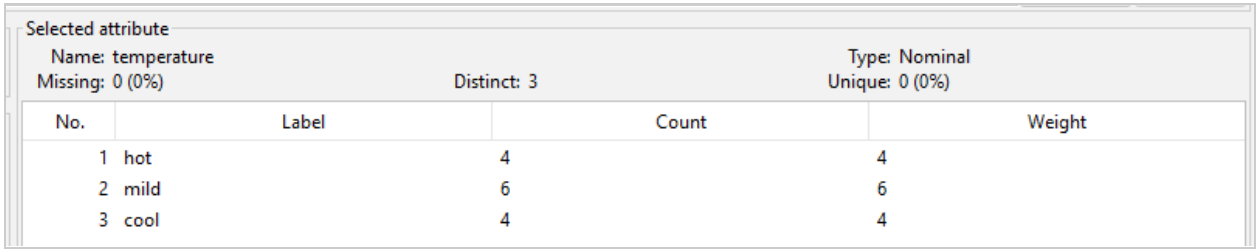
Name: play Type: Nominal
Missing: 0 (0%) Distinct: 2 Unique: 0 (0%)

No.	Label	Count	Weight
1	yes	9	9
2	no	5	5

(d) In the histogram on the bottom-right, which attributes are plotted on the X,Y-axes?How do You change the attributes plotted on the X,Y-axes?



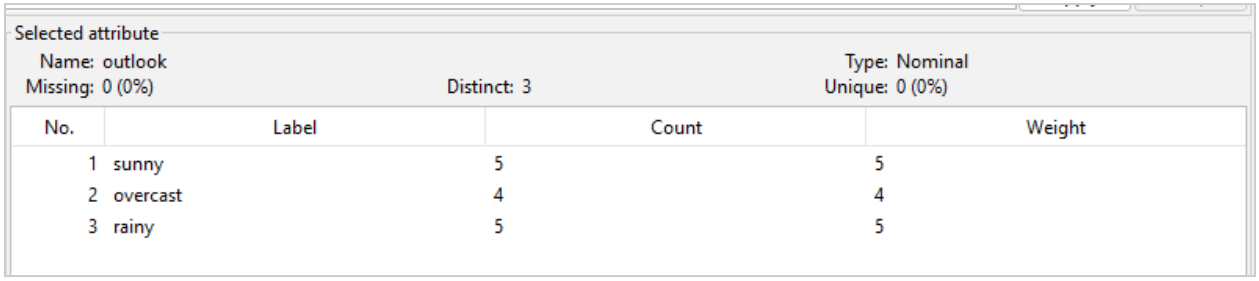
(e) How will you determine how many instances of each class are present in the data



Selected attribute

Name: temperature Type: Nominal
Missing: 0 (0%) Distinct: 3 Unique: 0 (0%)

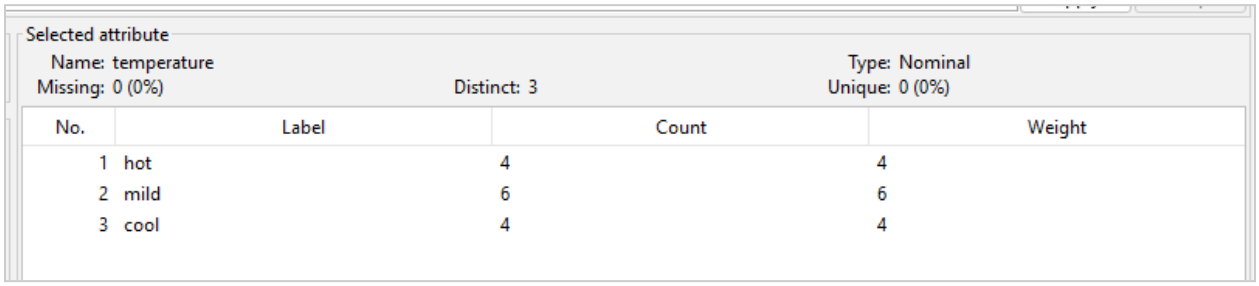
No.	Label	Count	Weight
1	hot	4	4
2	mild	6	6
3	cool	4	4



Selected attribute

Name: outlook Type: Nominal
Missing: 0 (0%) Distinct: 3 Unique: 0 (0%)

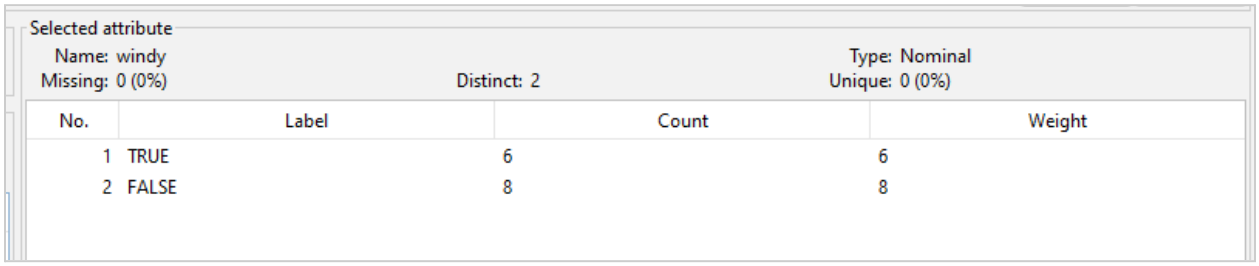
No.	Label	Count	Weight
1	sunny	5	5
2	overcast	4	4
3	rainy	5	5



Selected attribute

Name: temperature Type: Nominal
Missing: 0 (0%) Distinct: 3 Unique: 0 (0%)

No.	Label	Count	Weight
1	hot	4	4
2	mild	6	6
3	cool	4	4

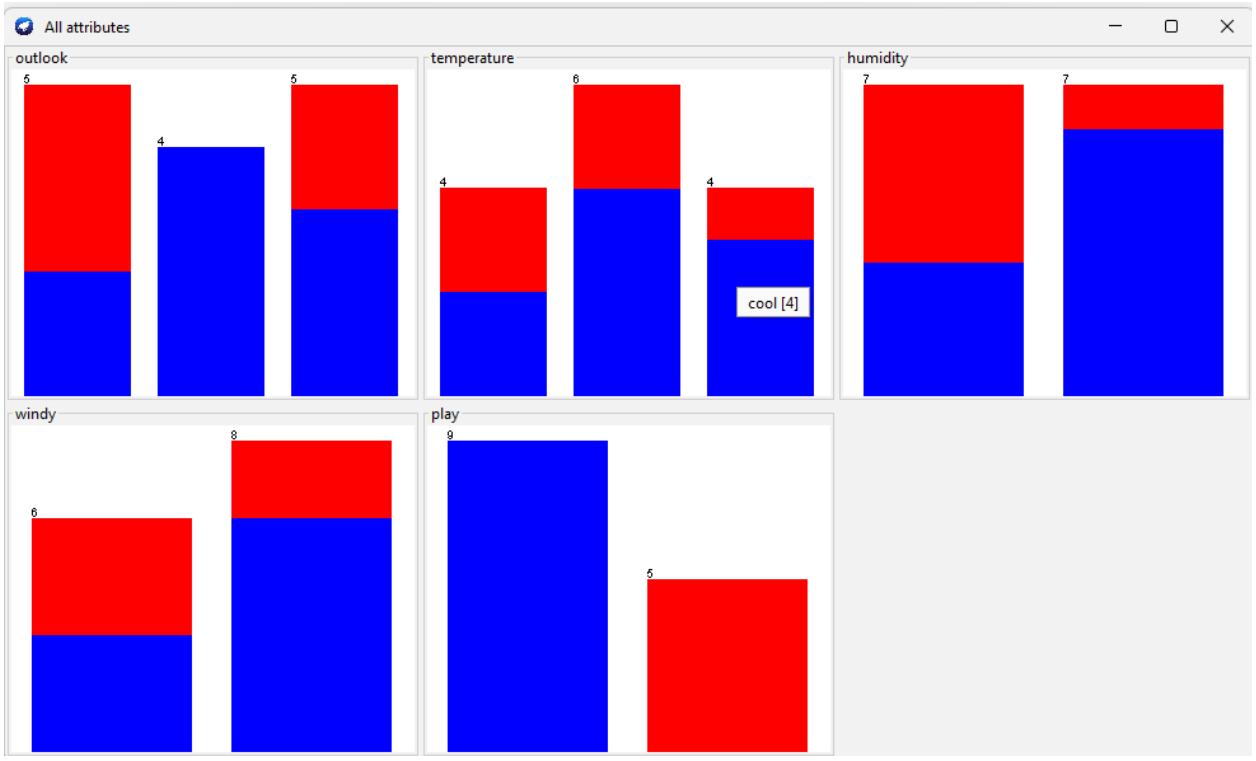


Selected attribute

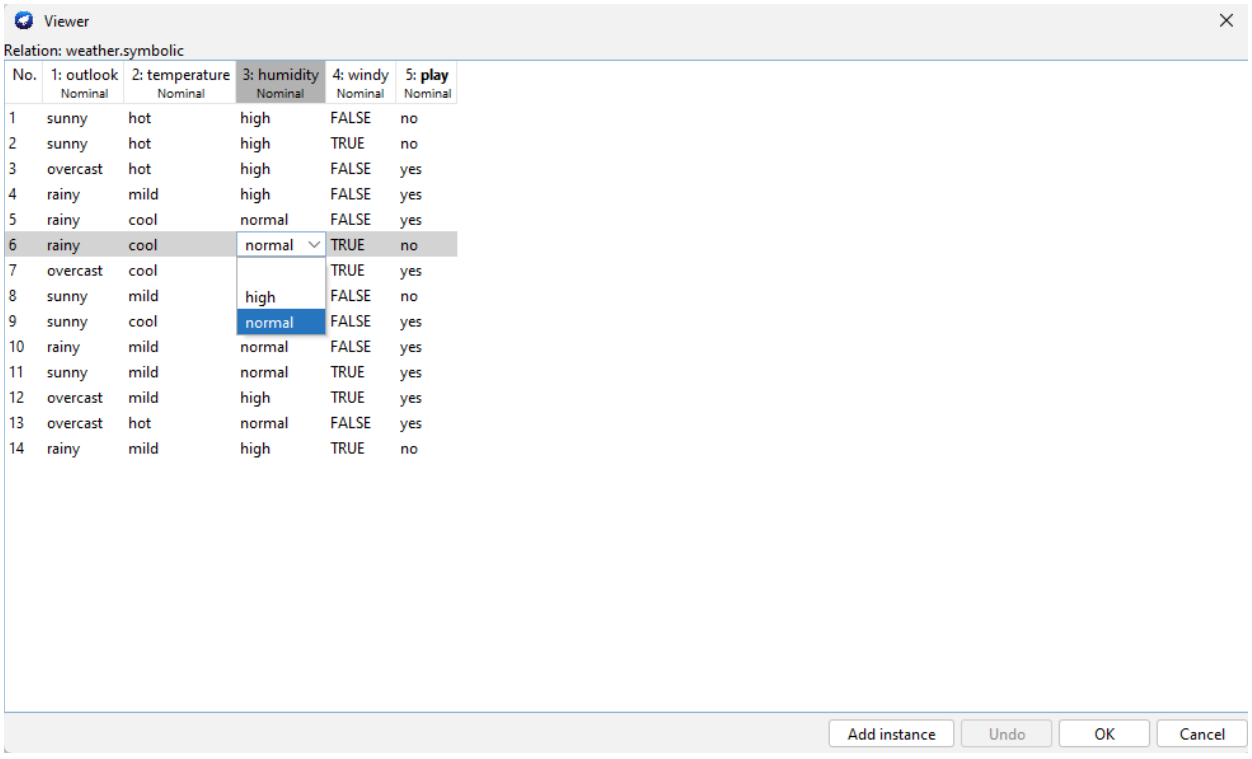
Name: windy Type: Nominal
Missing: 0 (0%) Distinct: 2 Unique: 0 (0%)

No.	Label	Count	Weight
1	TRUE	6	6
2	FALSE	8	8

(f) What happens when the Visualize All button is pressed?



(g) How will you view the instances in the dataset? How will you save the changes?



No.	1: outlook Nominal	2: temperature Nominal	3: humidity Nominal	4: windy Nominal	5: play Nominal
1	sunny	hot	high	FALSE	no
2	sunny	hot	high	TRUE	no
3	overcast	hot	high	FALSE	yes
4	rainy	mild	high	FALSE	yes
5	rainy	cool	normal	FALSE	yes
6	rainy	cool	normal	TRUE	no
7	overcast	cool		TRUE	yes
8	sunny	mild	high	FALSE	no
9	sunny	cool	normal	FALSE	yes
10	rainy	mild	normal	FALSE	yes
11	sunny	mild	normal	TRUE	yes
12	overcast	mild	high	TRUE	yes
13	overcast	hot	normal	FALSE	yes
14	rainy	mild	high	TRUE	no

2. Load the weather dataset and perform the following tasks:

(a) Use the unsupervised filter RemoveWithValues to remove all instances where the attribute 'humidity' has the value 'high'?

Viewer

Relation: weather.symbolic-weka.filters.unsupervised.instance.RemoveWithValues-S0.0-C3-Lfirst

No.	1: outlook Nominal	2: temperature Nominal	3: **humidity** Nominal	4: windy Nominal	5: play Nominal
1	rainy	cool	normal	FALSE	yes
2	rainy	cool	normal	TRUE	no
3	overcast	cool	normal	TRUE	yes
4	sunny	cool	normal	FALSE	yes
5	rainy	mild	normal	FALSE	yes
6	sunny	mild	normal	TRUE	yes
7	overcast	hot	normal	FALSE	yes

(b) Undo the effect of the filter.

Viewer

Relation: weather.symbolic

No.	1: outlook Nominal	2: temperature Nominal	3: humidity Nominal	4: windy Nominal	5: **play** Nominal
1	sunny	hot	high	FALSE	no
2	sunny	hot	high	TRUE	no
3	overcast	hot	high	FALSE	yes
4	rainy	mild	high	FALSE	yes
5	rainy	cool	normal	FALSE	yes
6	rainy	cool	normal	TRUE	no
7	overcast	cool	normal	TRUE	yes
8	sunny	mild	high	FALSE	no
9	sunny	cool	normal	FALSE	yes
10	rainy	mild	normal	FALSE	yes
11	sunny	mild	normal	TRUE	yes
12	overcast	mild	high	TRUE	yes
13	overcast	hot	normal	FALSE	yes
14	rainy	mild	high	TRUE	no

Add instance | Undo | OK | Cancel

(c) Answer the following questions:

(i) What is meant by filtering in Weka?
Ans: In Weka, a filter is a tool used to preprocess and manipulate data before it is used to train a machine learning model. Filters can be used to perform various operations on the data such as removing attributes, transforming values, normalizing data, or selecting subsets of data based on certain criteria.

Filters in Weka can be applied to the entire dataset or a subset of the dataset. They can be applied before or after splitting the dataset into training and testing sets. Filters are an important tool for preparing data for machine learning models and can help improve the accuracy of the models.

(ii) Which panel is used for filtering a dataset?
Ans: In Weka, the panel used for filtering a dataset is called the "Preprocess" panel. It is one of the main panels in the Weka Explorer interface, which is the graphical user interface for Weka. In the Preprocess panel, you can select the filter you want to apply from a list of available filters, configure its parameters, and apply it to the dataset. Once the filter is applied, you can save the filtered dataset for further analysis or use it directly for training and testing machine learning models.

(iii) What are the two main types of filters in Weka?
Ans: In Weka, the two main types of filters are supervised and unsupervised filters.

Supervised filters: These filters use information from the class labels (i.e., the target variable) to modify the data. They are used when the class labels are available and the filter modifies the data based on the class labels. For example, the AttributeSelection filter can be used to select a subset of the most informative attributes for a given classification task.

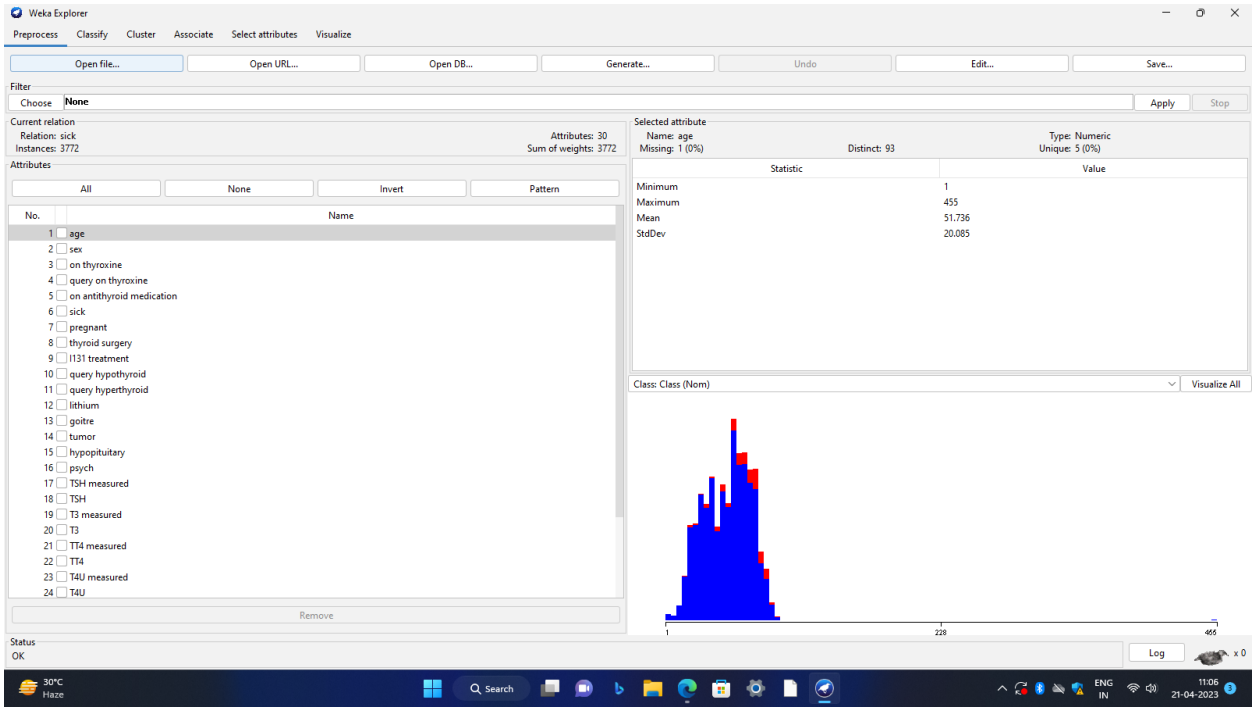
Unsupervised filters: These filters do not use any information from the class labels and modify the data based only on its attributes. They are used when the class labels are not available or when the filter should not depend on the class labels. For example, the PrincipalComponents filter can be used to reduce the dimensionality of the data by transforming the attributes into a smaller set of uncorrelated variables.

(iv) What is the difference between the two types of filters? What is the difference between and attribute filter and an instance filter?
Ans: An attribute filter in Weka is a type of filter that modifies the attributes (columns) of the dataset. For example, an attribute filter may remove certain attributes that are not relevant to the analysis, transform the values of an attribute to a different scale, or discretize a numeric attribute into categorical values. Attribute filters are applied to all instances (rows) of the dataset, and the same filter is applied to all attributes.

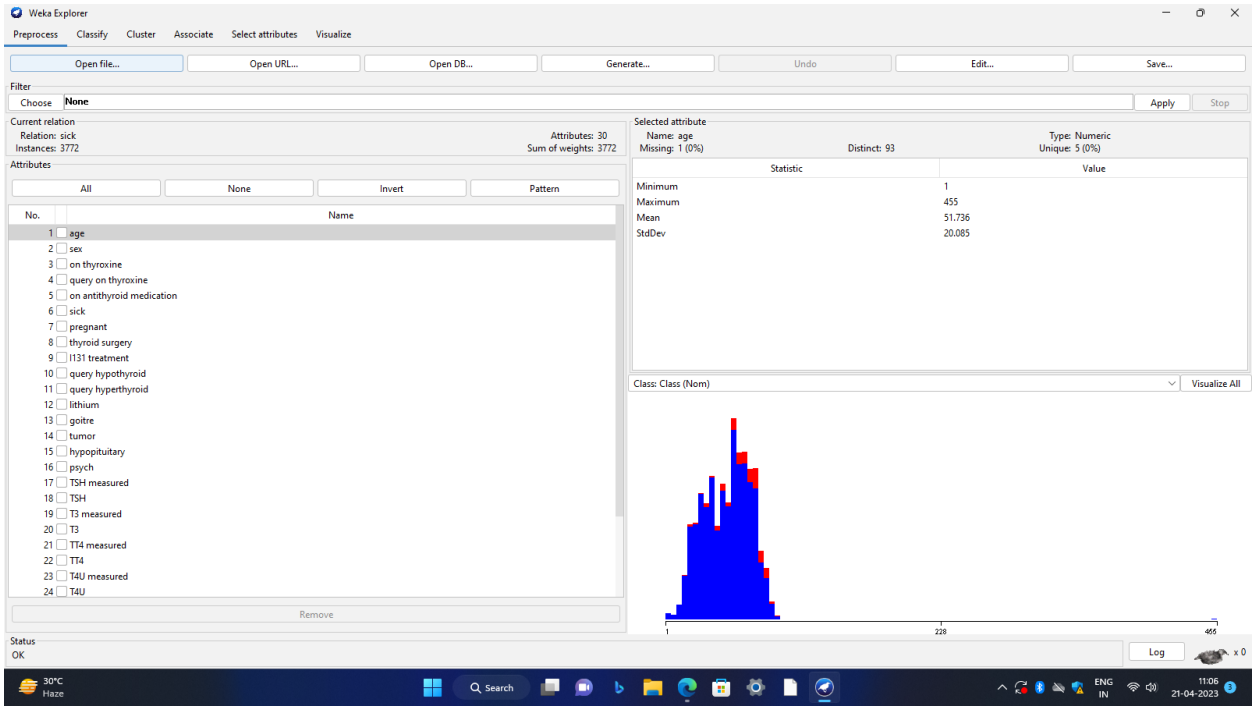
On the other hand, an instance filter in Weka is a type of filter that modifies the instances (rows) of the dataset. For example, an instance filter may remove certain instances that are outliers or duplicates, balance the class distribution of the dataset, or resample the dataset to reduce its size. Instance filters are applied to all attributes of the dataset, and the same filter is applied to all instances.

Part I

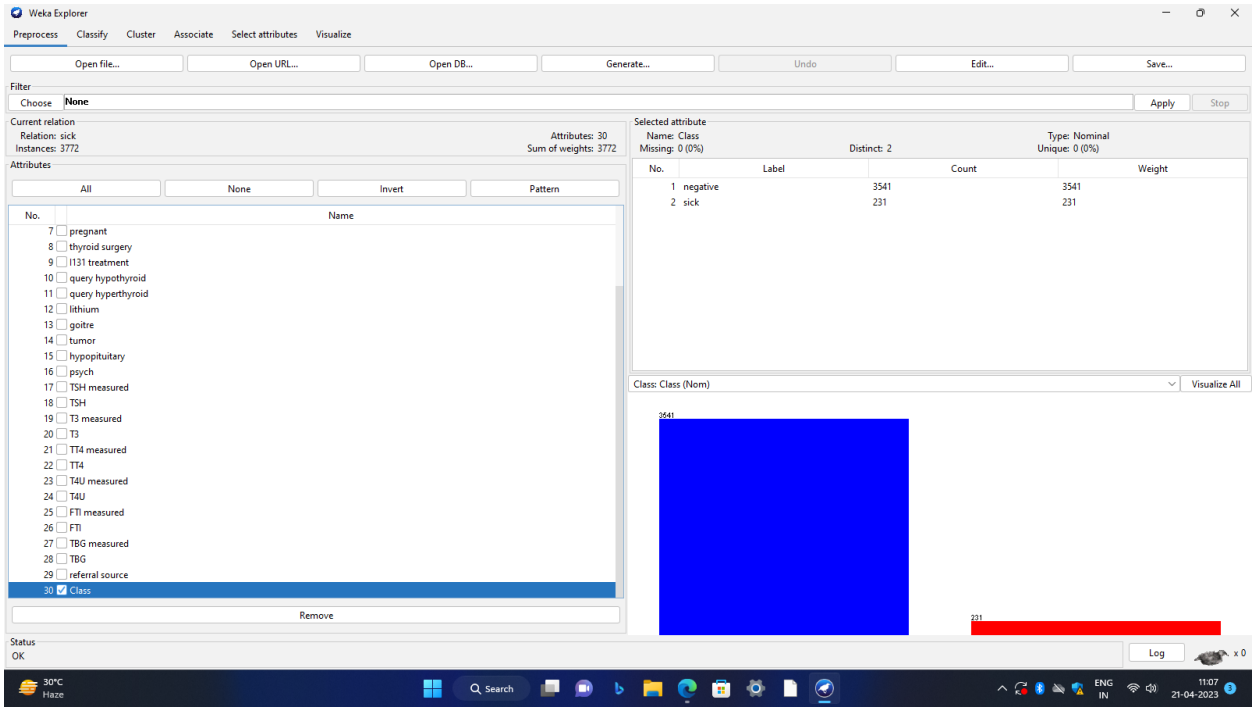
1.2



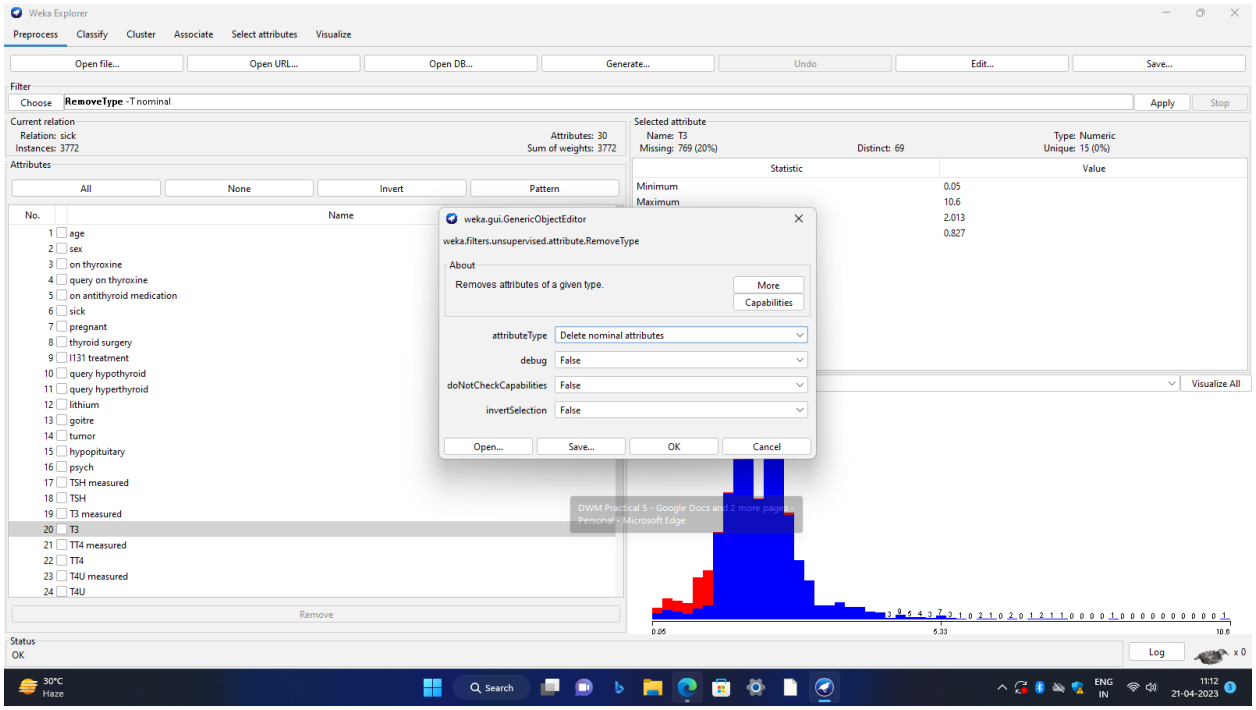
1.3

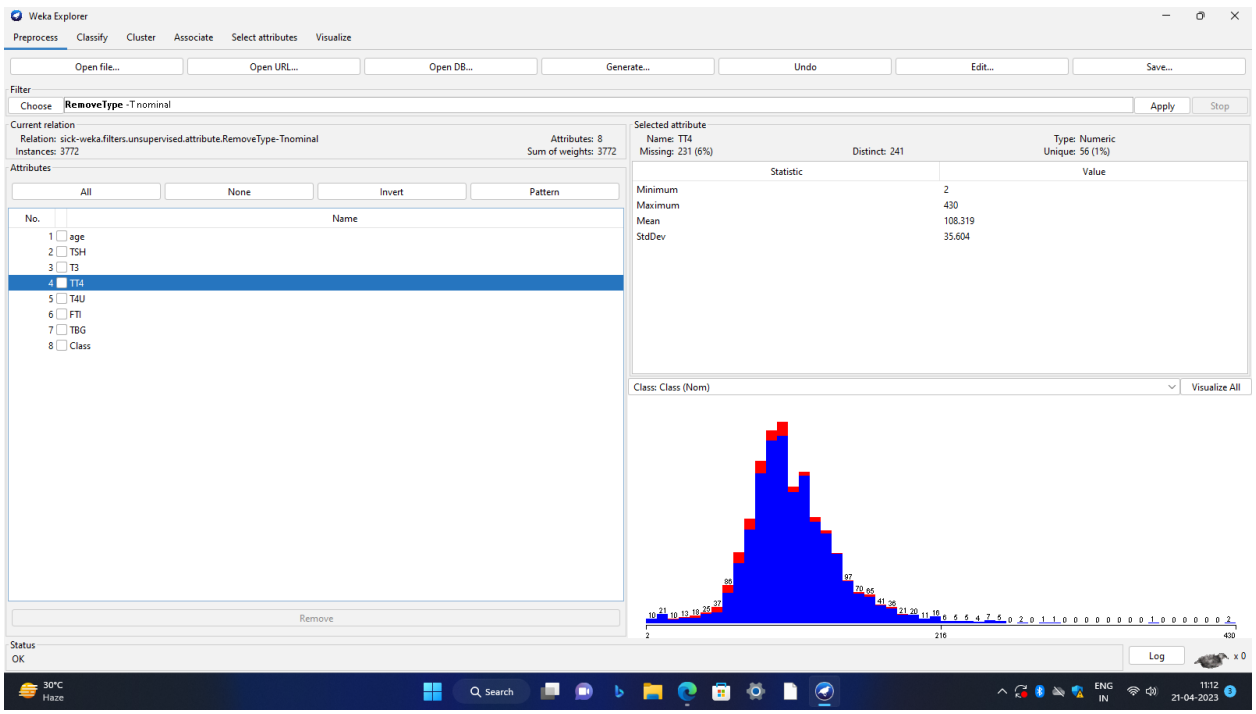

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Open file...
Open URL...
Open DB...
Generate...
Undo
Edit...
Save...
Filter
Choose
None
Apply
Stop
Current relation
Relation: sick
Instances: 3772
Attributes: 30
Sum of weights: 3772
Attributes
All
None
Invert
Pattern
No.
Name
1 age
2 sex
3 on thyroxine
4 query on thyroxine
5 on antithyroid medication
6 sick
7 pregnant
8 thyroid surgery
9 I131 treatment
10 query hypothyroid
11 query hyperthyroid
12 lithium
13 goitre
14 tumor
15 hypopituitary
16 psych
17 TSH measured
18 TSH
19 T3 measured
20 T3
21 TT4 measured
22 TT4
23 T4U measured
24 T4U
Remove
Selected attribute
Name: age
Missing: 1 (0%)
Distinct: 93
Type: Numeric
Unique: 5 (0%)
Statistic
Value
Minimum 1
Maximum 455
Mean 51.736
StdDev 20.085
Class: Class (Nom)
Visualize All
1
228
455
Status
OK
Log
x 0
30°C
Haze
Search
ENG
IN
11:06
21-04-2023

1.4

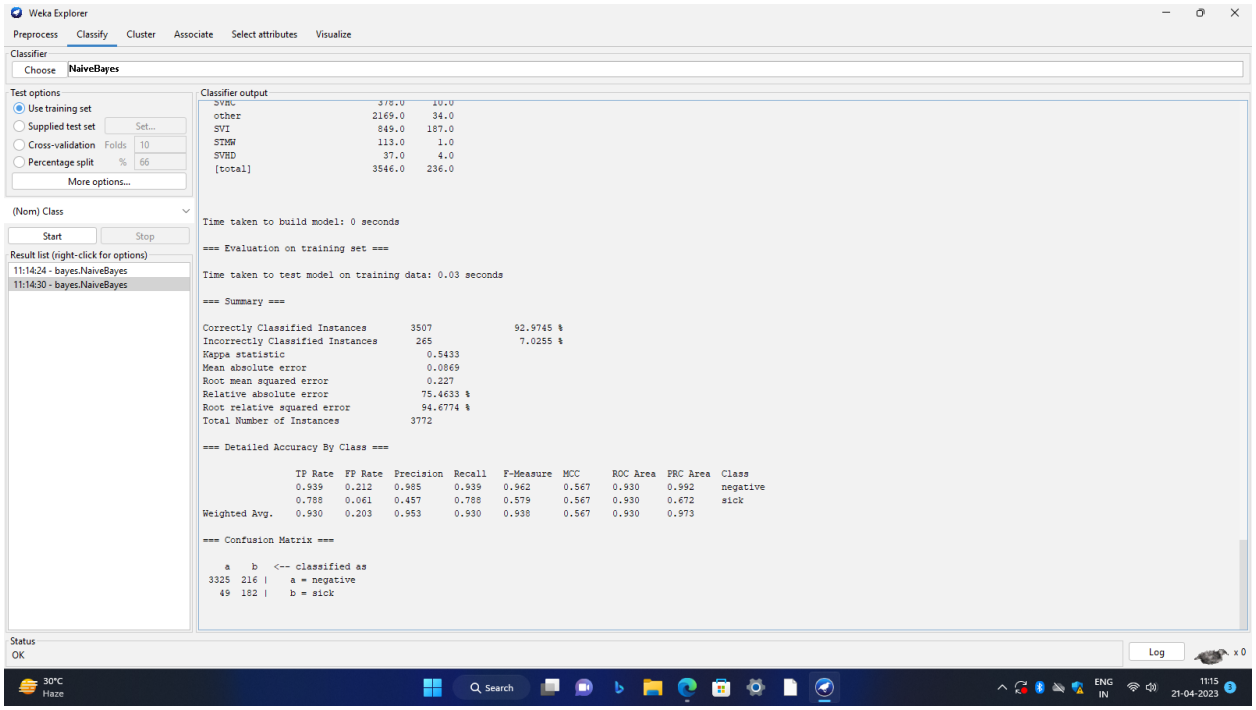

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Open file...
Open URL...
Open DB...
Generate...
Undo
Edit...
Save...
Filter
Choose
None
Apply
Stop
Current relation
Relation: sick
Instances: 3772
Attributes: 30
Sum of weights: 3772
Attributes
All
None
Invert
Pattern
No.
Name
7 pregnant
8 thyroid surgery
9 I131 treatment
10 query hypothyroid
11 query hyperthyroid
12 lithium
13 goitre
14 tumor
15 hypopituitary
16 psych
17 TSH measured
18 TSH
19 T3 measured
20 T3
21 TT4 measured
22 TT4
23 T4U measured
24 T4U
25 FTI measured
26 FTI
27 TBG measured
28 TBG
29 referral source
30 Class
Remove
Selected attribute
Name: Class
Missing: 0 (0%)
Distinct: 2
Type: Nominal
Unique: 0 (0%)
No.
Label
Count
Weight
1 negative 3541 3541
2 sick 231 231
Class: Class (Nom)
Visualize All
3541
231
Status
OK
Log
x 0
30°C
Haze
Search
ENG
IN
11:07
21-04-2023

1.5

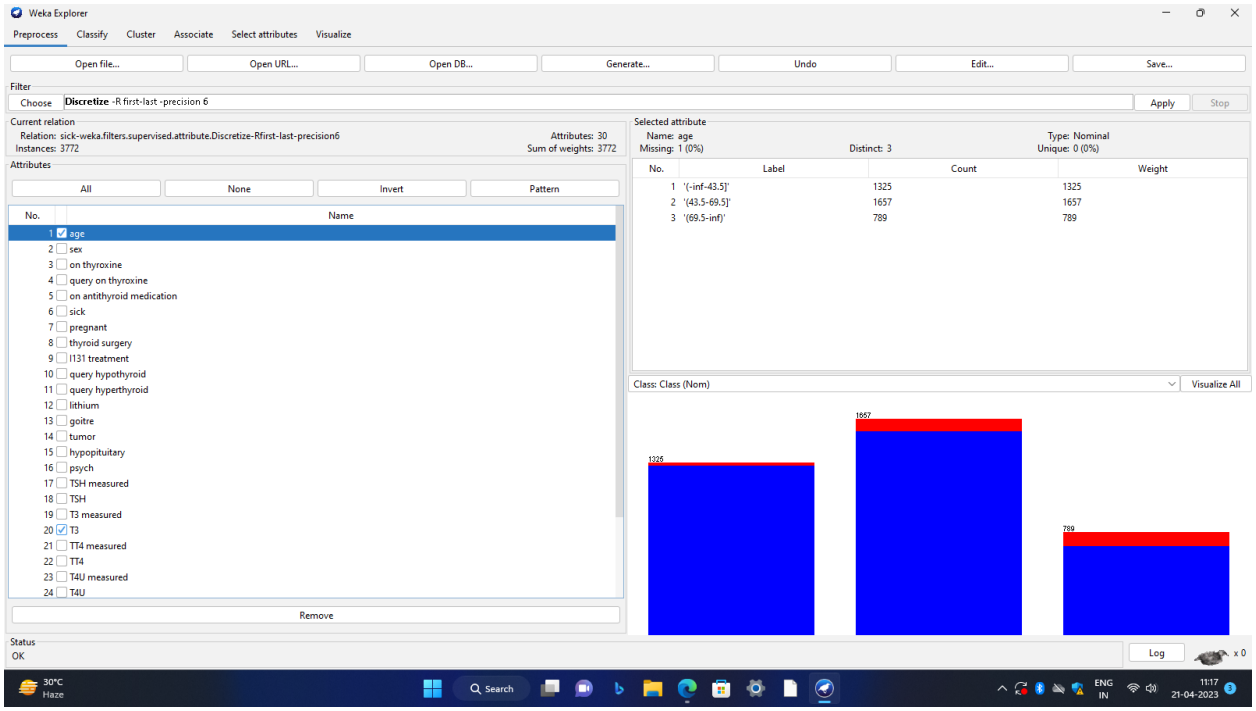

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Open file...
Open URL...
Open DB...
Generate...
Undo
Edit...
Save...
Filter
Choose
RemoveType -T nominal
Apply
Stop
Current relation
Relation: sick
Instances: 3772
Attributes: 30
Sum of weights: 3772
Attributes
All
None
Invert
Pattern
No.
Name
1 age
2 sex
3 on thyroxine
4 query on thyroxine
5 on antithyroid medication
6 sick
7 pregnant
8 thyroid surgery
9 I131 treatment
10 query hypothyroid
11 query hyperthyroid
12 lithium
13 goitre
14 tumor
15 hypopituitary
16 psych
17 TSH measured
18 TSH
19 T3 measured
20 T3
21 TT4 measured
22 TT4
23 T4U measured
24 T4U
Remove
Selected attribute
Name: T3
Missing: 769 (20%)
Distinct: 69
Type: Numeric
Unique: 15 (0%)
Statistic
Value
Minimum 0.05
Maximum 10.6
2.013
0.827
Visualize All
weka.gui.GenericObjectEditor
weka.filters.unsupervised.attribute.RemoveType
About
Removes attributes of a given type.
More
Capabilities
attributeType
Delete nominal attributes
debug
False
doNotCheckCapabilities
False
invertSelection
False
Open...
Save...
OK
Cancel
DWM Practical 5 - Google Docs and 2 more pages - Personal - Microsoft Edge
0.05
5.33
10.6
Status
OK
Log
x 0
30°C
Haze
Search
ENG
IN
11:12
21-04-2023

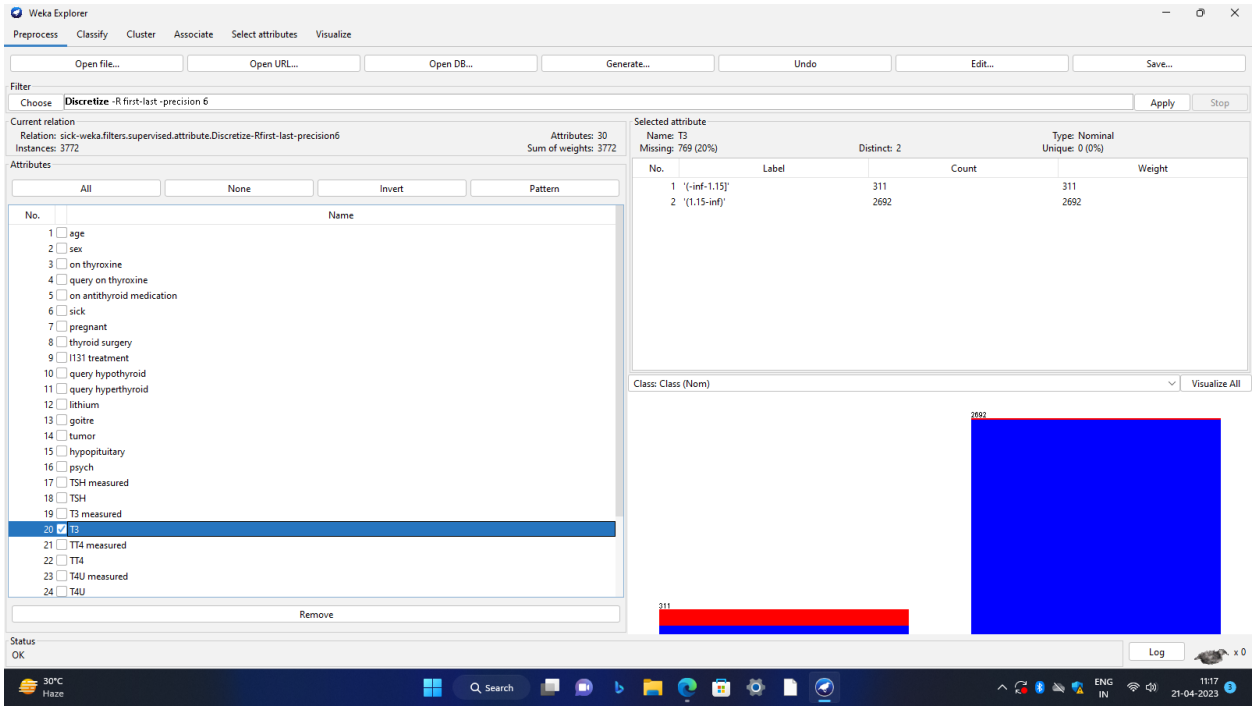

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Open file...
Open URL...
Open DB...
Generate...
Undo
Edit...
Save...
Filter
Choose
RemoveType -T nominal
Apply
Stop
Current relation
Relation: sick-weka.filters.unsupervised.attribute.RemoveType-Tnominal
Instances: 3772
Attributes: 8
Sum of weights: 3772
Attributes
All
None
Invert
Pattern
No.
Name
1 age
2 TSH
3 T3
4 TT4
5 T4U
6 FTI
7 TBG
8 Class
Remove
Selected attribute
Name: TT4
Missing: 231 (6%)
Distinct: 241
Type: Numeric
Unique: 56 (1%)
Statistic
Value
Minimum 2
Maximum 430
Mean 108.319
StdDev 35.604
Class: Class (Nom)
Visualize All
2
216
430
Status
OK
Log
x 0
30°C
Haze
Search
ENG
IN
11:12
21-04-2023

1.6


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)
11:14:24 - bayes.NaiveBayes
11:14:30 - bayes.NaiveBayes
Classifier output
SVMC 378.0 10.0
other 2169.0 34.0
SVI 849.0 187.0
STMH 113.0 1.0
SVHD 37.0 4.0
[total] 3546.0 236.0
Time taken to build model: 0 seconds
=== Evaluation on training set ===
Time taken to test model on training data: 0.03 seconds
=== Summary ===
Correctly Classified Instances 3507 92.9745 %
Incorrectly Classified Instances 265 7.0255 %
Kappa statistic 0.5433
Mean absolute error 0.0869
Root mean squared error 0.227
Relative absolute error 75.4633 %
Root relative squared error 94.6774 %
Total Number of Instances 3772
=== Detailed Accuracy By Class ===
TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
0.939 0.212 0.985 0.939 0.962 0.567 0.930 0.992 negative
0.788 0.061 0.457 0.788 0.579 0.567 0.930 0.672 sick
Weighted Avg. 0.930 0.203 0.953 0.930 0.938 0.567 0.930 0.973
=== Confusion Matrix ===
a b <-- classified as
3325 216 | a = negative
49 182 | b = sick
Status
OK
Log
x 0
30°C Haze
Search
ENG IN
11:15 21-04-2023

2.2

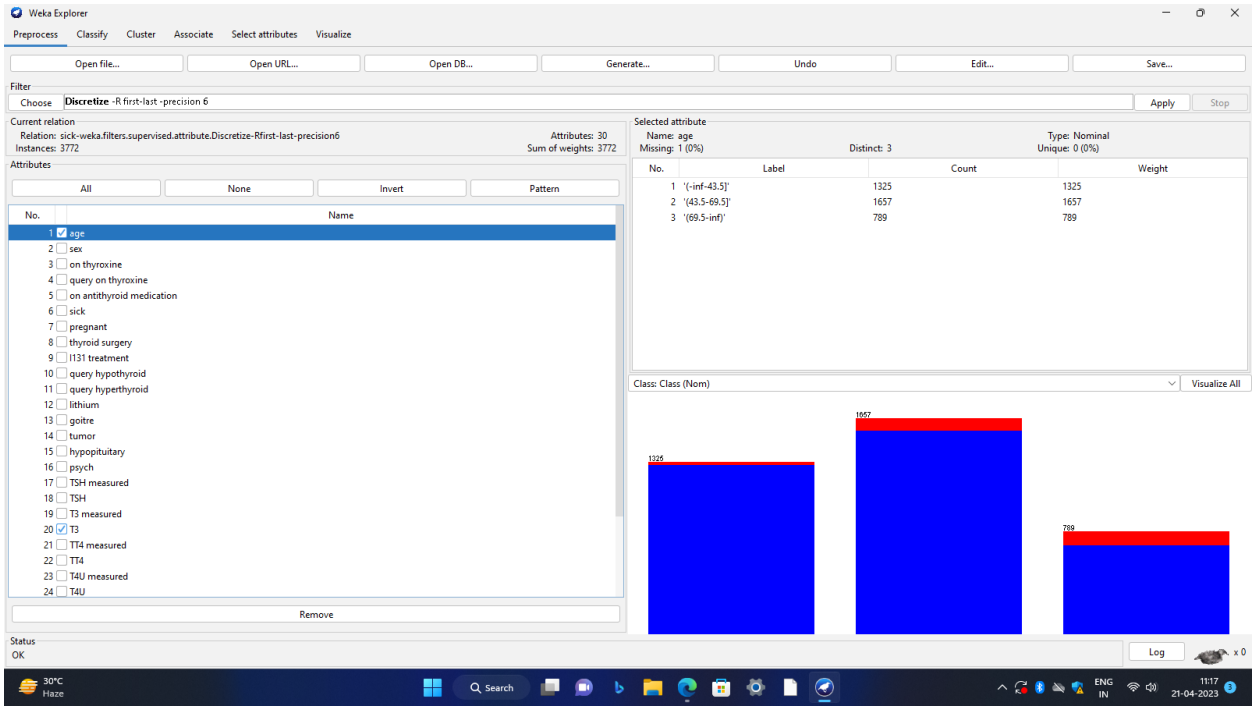

Weka Explorer
Preprocess Classify Cluster Associate Select attributes Visualize
Open file... Open URL... Open DB... Generate... Undo Edit... Save...
Filter
Choose Discretize -R first-last -precision 6
Apply Stop
Current relation
Relation: sick-weka.filters.supervised.attribute.Discretize-Rfirst-last-precision6
Instances: 3772
Attributes: 30
Sum of weights: 3772
Attributes
All None Invert Pattern
No. Name
1 age
2 sex
3 on thyroxine
4 query on thyroxine
5 on antithyroid medication
6 sick
7 pregnant
8 thyroid surgery
9 I131 treatment
10 query hypothyroid
11 query hyperthyroid
12 lithium
13 goitre
14 tumor
15 hypopituitary
16 psych
17 TSH measured
18 TSH
19 T3 measured
20 T3
21 TT4 measured
22 TT4
23 T4U measured
24 T4U
Remove
Selected attribute
Name: age
Missing: 1 (0%)
Distinct: 3
Type: Nominal
Unique: 0 (0%)
No. Label Count Weight
1 '(-inf-43.5]' 1325 1325
2 '(43.5-69.5]' 1657 1657
3 '(69.5-inf)' 789 789
Class: Class (Nom)
Visualize All
1325
1657
789
Status
OK
Log
x 0
30°C Haze
Search
ENG IN
11:17 21-04-2023



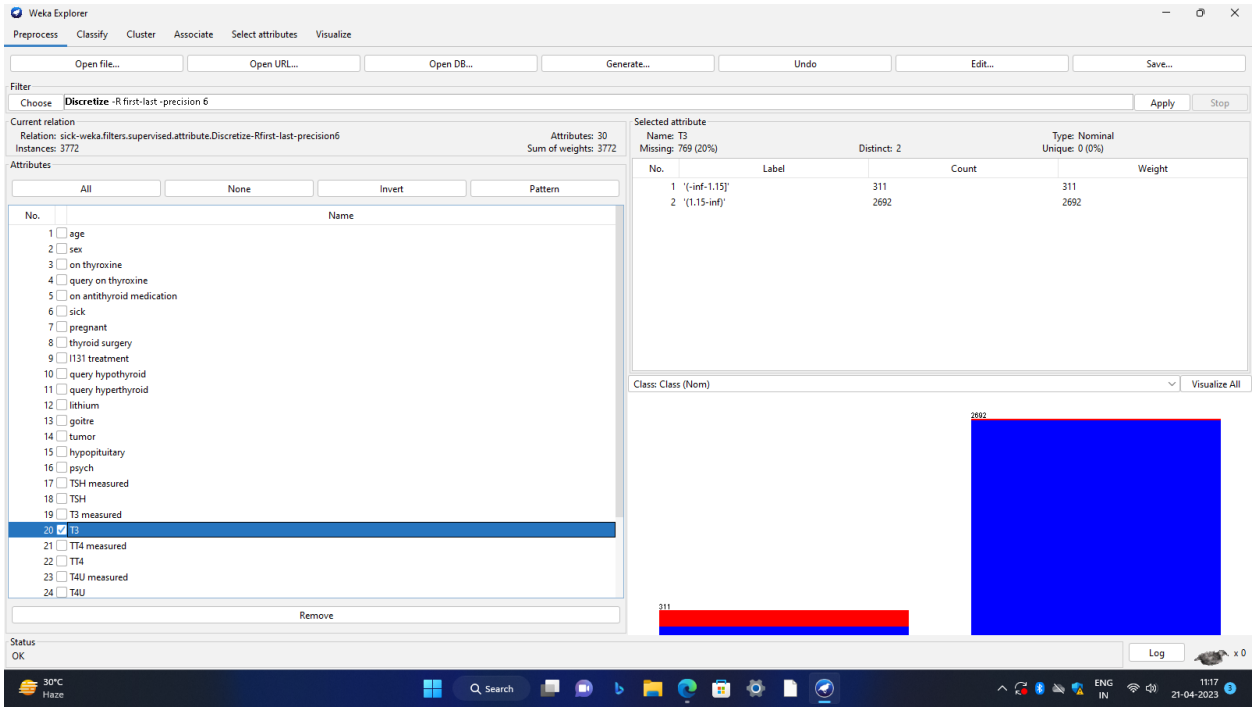
2.3

NUmeric data is converted to nominal data

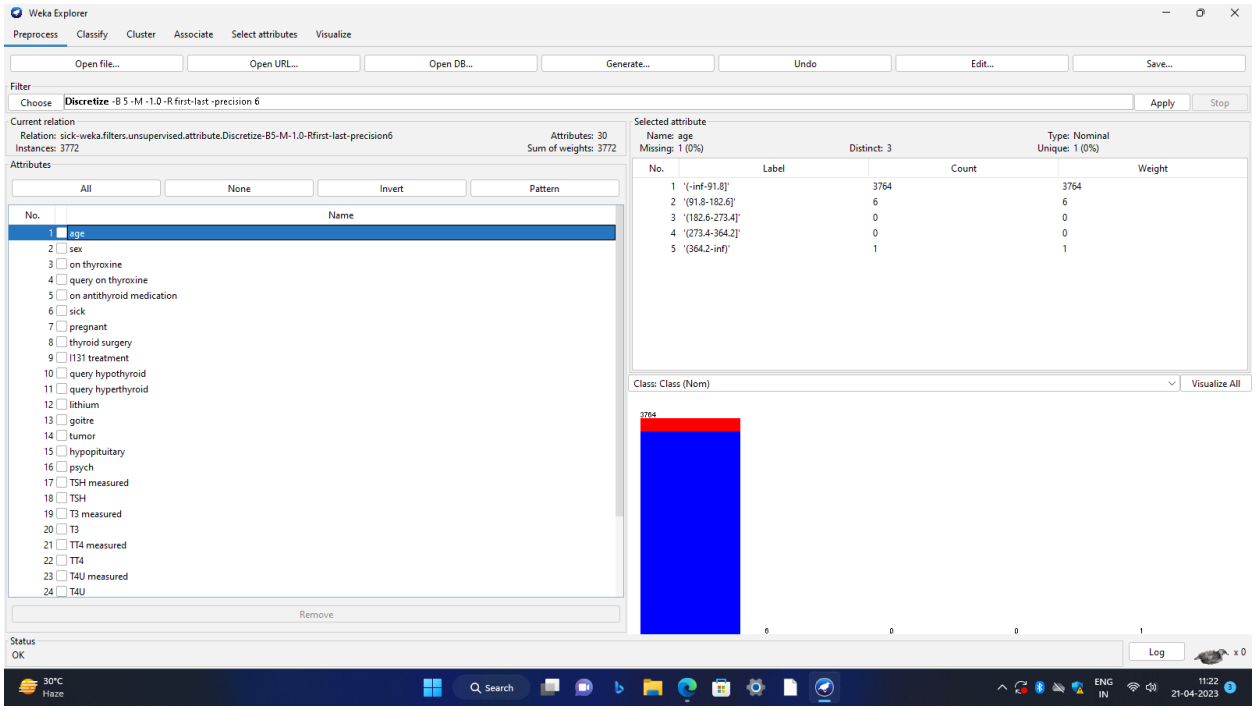
2.4


Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Open file...
Open URL...
Open DB...
Generate...
Undo
Edit...
Save...
Filter
Choose
Discretize -R first-last -precision 6
Apply
Stop
Current relation
Relation: sick-weka.filters.supervised.attribute.Discretize-Rfirst-last-precision6
Instances: 3772
Attributes: 30
Sum of weights: 3772
Attributes
All
None
Invert
Pattern
No.
Name
1 age
2 sex
3 on thyroxine
4 query on thyroxine
5 on antithyroid medication
6 sick
7 pregnant
8 thyroid surgery
9 I131 treatment
10 query hypothyroid
11 query hyperthyroid
12 lithium
13 goitre
14 tumor
15 hypopituitary
16 psych
17 TSH measured
18 TSH
19 T3 measured
20 T3
21 TT4 measured
22 TT4
23 T4U measured
24 T4U
Remove
Selected attribute
Name: age
Missing: 1 (0%)
Distinct: 3
Type: Nominal
Unique: 0 (0%)
No.
Label
Count
Weight
1 '(-inf-43.5]' 1325 1325
2 '(43.5-69.5]' 1657 1657
3 '(69.5-inf)' 789 789
Class: Class (Nom)
Visualize All
1325
1657
789
Status
OK
Log
x 0
30°C
Haze
Search
ENG
IN
11:17
21-04-2023

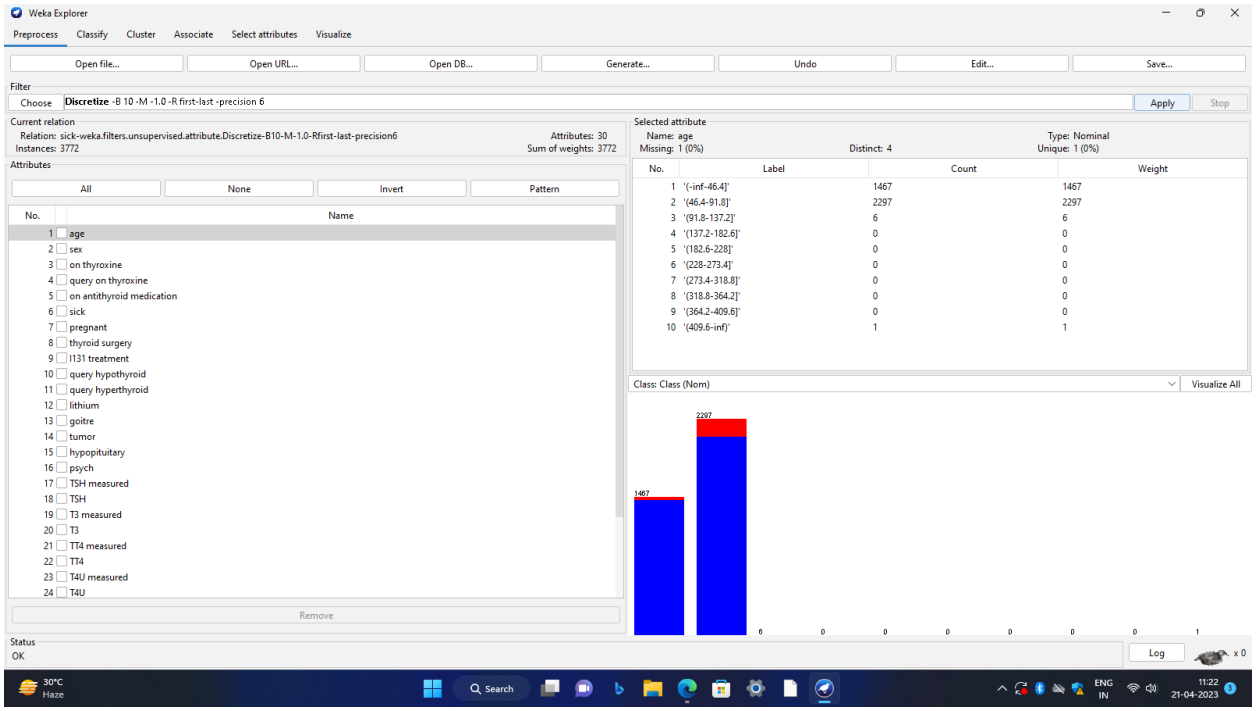
,


Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Open file...
Open URL...
Open DB...
Generate...
Undo
Edit...
Save...
Filter
Choose
Discretize -R first-last -precision 6
Apply
Stop
Current relation
Relation: sick-weka.filters.supervised.attribute.Discretize-Rfirst-last-precision6
Instances: 3772
Attributes: 30
Sum of weights: 3772
Attributes
All
None
Invert
Pattern
No.
Name
1 age
2 sex
3 on thyroxine
4 query on thyroxine
5 on antithyroid medication
6 sick
7 pregnant
8 thyroid surgery
9 I131 treatment
10 query hypothyroid
11 query hyperthyroid
12 lithium
13 goitre
14 tumor
15 hypopituitary
16 psych
17 TSH measured
18 TSH
19 T3 measured
20 T3
21 TT4 measured
22 TT4
23 T4U measured
24 T4U
Remove
Selected attribute
Name: T3
Missing: 769 (20%)
Distinct: 2
Type: Nominal
Unique: 0 (0%)
No.
Label
Count
Weight
1 '(-inf-1.15]' 311 311
2 '(1.15-inf)' 2692 2692
Class: Class (Nom)
Visualize All
2692
311
Status
OK
Log
x 0
30°C
Haze
Search
ENG
IN
11:17
21-04-2023

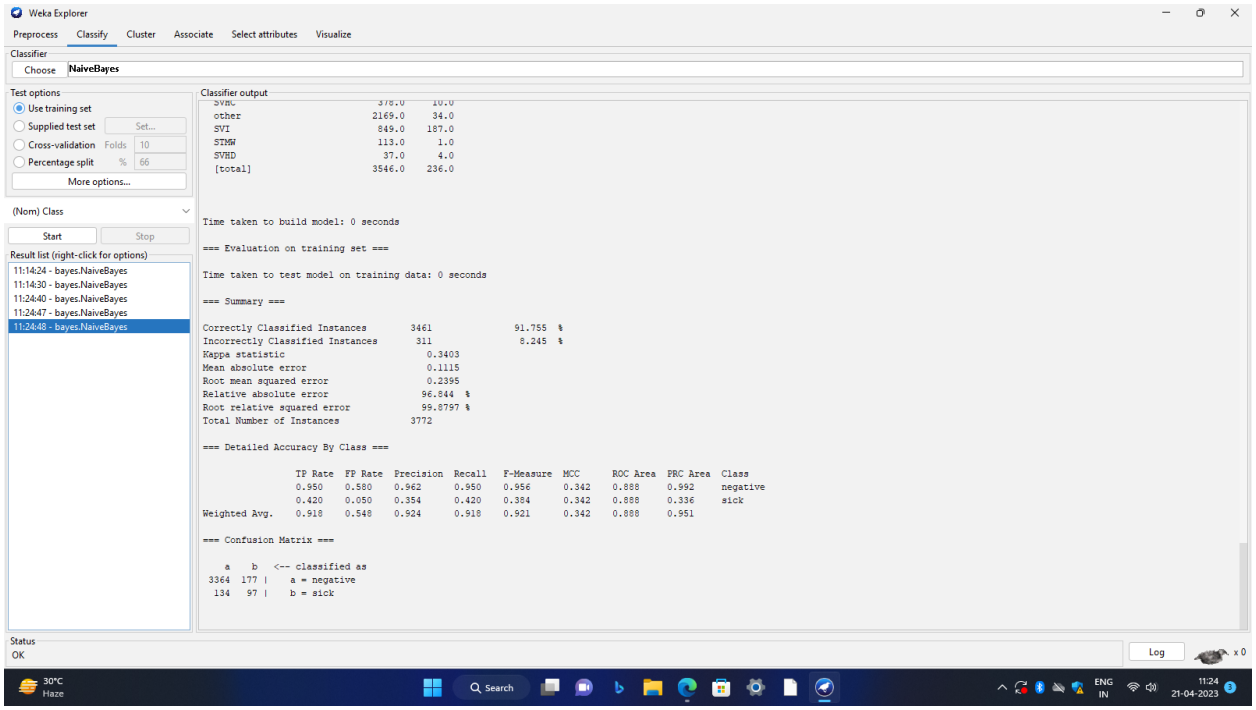
2.6.1


Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Open file...
Open URL...
Open DB...
Generate...
Undo
Edit...
Save...
Filter
Choose
Discretize -B 5 -M -1.0 -R first-last -precision 6
Apply
Stop
Current relation
Relation: sick-weka.filters.unsupervised.attribute.Discretize-B5-M-1.0-Rfirst-last-precision6
Instances: 3772
Attributes: 30
Sum of weights: 3772
Attributes
All
None
Invert
Pattern
No.
Name
1 age
2 sex
3 on thyroxine
4 query on thyroxine
5 on antithyroid medication
6 sick
7 pregnant
8 thyroid surgery
9 I131 treatment
10 query hypothyroid
11 query hyperthyroid
12 lithium
13 goitre
14 tumor
15 hypopituitary
16 psych
17 TSH measured
18 TSH
19 T3 measured
20 T3
21 TT4 measured
22 TT4
23 T4U measured
24 T4U
Remove
Selected attribute
Name: age
Missing: 1 (0%)
Distinct: 3
Type: Nominal
Unique: 1 (0%)
No.
Label
Count
Weight
1 '(-inf-91.8]' 3764 3764
2 '(91.8-182.6]' 6 6
3 '(182.6-273.4]' 0 0
4 '(273.4-364.2]' 0 0
5 '(364.2-inf)' 1 1
Class: Class (Nom)
Visualize All
3764
Status
OK
Log
x 0
30°C
Haze
Search
ENG
IN
11:22
21-04-2023

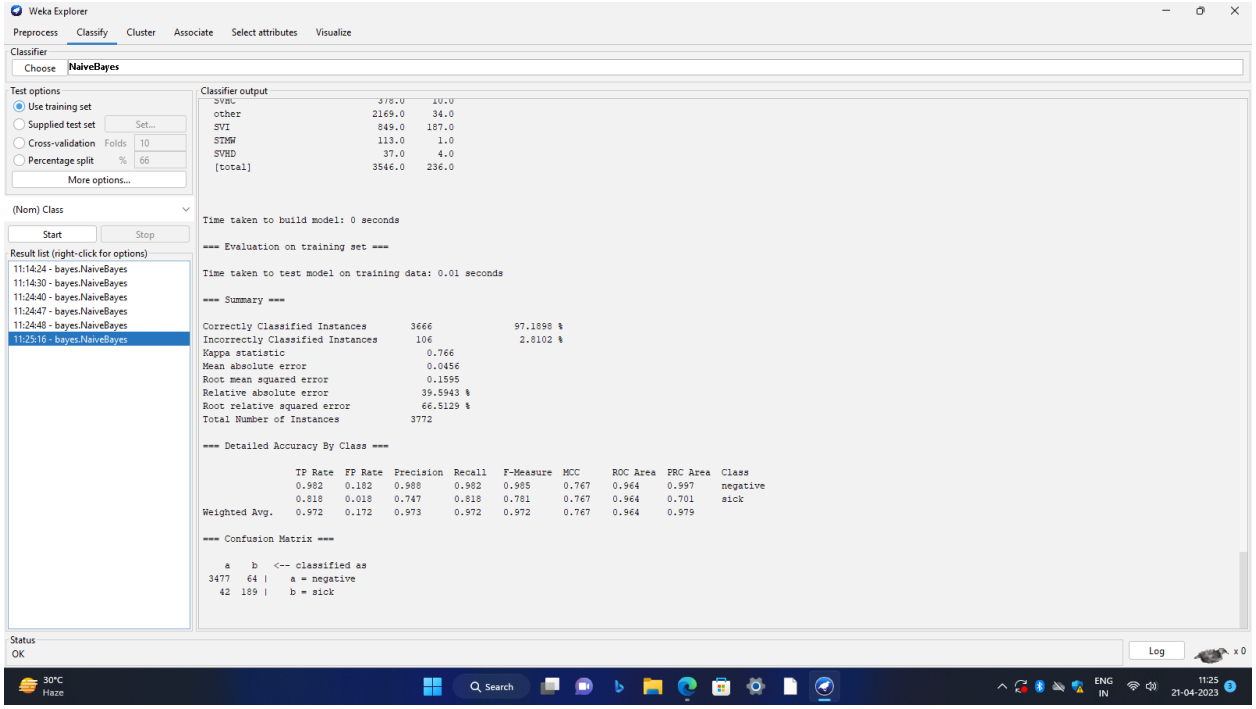
2.6.2


Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Open file...
Open URL...
Open DB...
Generate...
Undo
Edit...
Save...
Filter
Choose
Discretize -B 10 -M -1.0 -R first-last -precision 6
Apply
Stop
Current relation
Relation: sick-weka.filters.unsupervised.attribute.Discretize-B10-M-1.0-Rfirst-last-precision6
Instances: 3772
Attributes: 30
Sum of weights: 3772
Attributes
All
None
Invert
Pattern
No.
Name
1 age
2 sex
3 on thyroxine
4 query on thyroxine
5 on antithyroid medication
6 sick
7 pregnant
8 thyroid surgery
9 I131 treatment
10 query hypothyroid
11 query hyperthyroid
12 lithium
13 goitre
14 tumor
15 hypopituitary
16 psych
17 TSH measured
18 TSH
19 T3 measured
20 T3
21 TT4 measured
22 TT4
23 T4U measured
24 T4U
Remove
Selected attribute
Name: age
Missing: 1 (0%)
Distinct: 4
Type: Nominal
Unique: 1 (0%)
No.
Label
Count
Weight
1 '(-inf-46.4]' 1467 1467
2 '(46.4-91.8]' 2297 2297
3 '(91.8-137.2]' 6 6
4 '(137.2-182.6]' 0 0
5 '(182.6-228]' 0 0
6 '(228-273.4]' 0 0
7 '(273.4-318.8]' 0 0
8 '(318.8-364.2]' 0 0
9 '(364.2-409.6]' 0 0
10 '(409.6-inf)' 1 1
Class: Class (Nom)
Visualize All
2297
1467
Status
OK
Log
x 0
30°C
Haze
Search
ENG
IN
11:22
21-04-2023

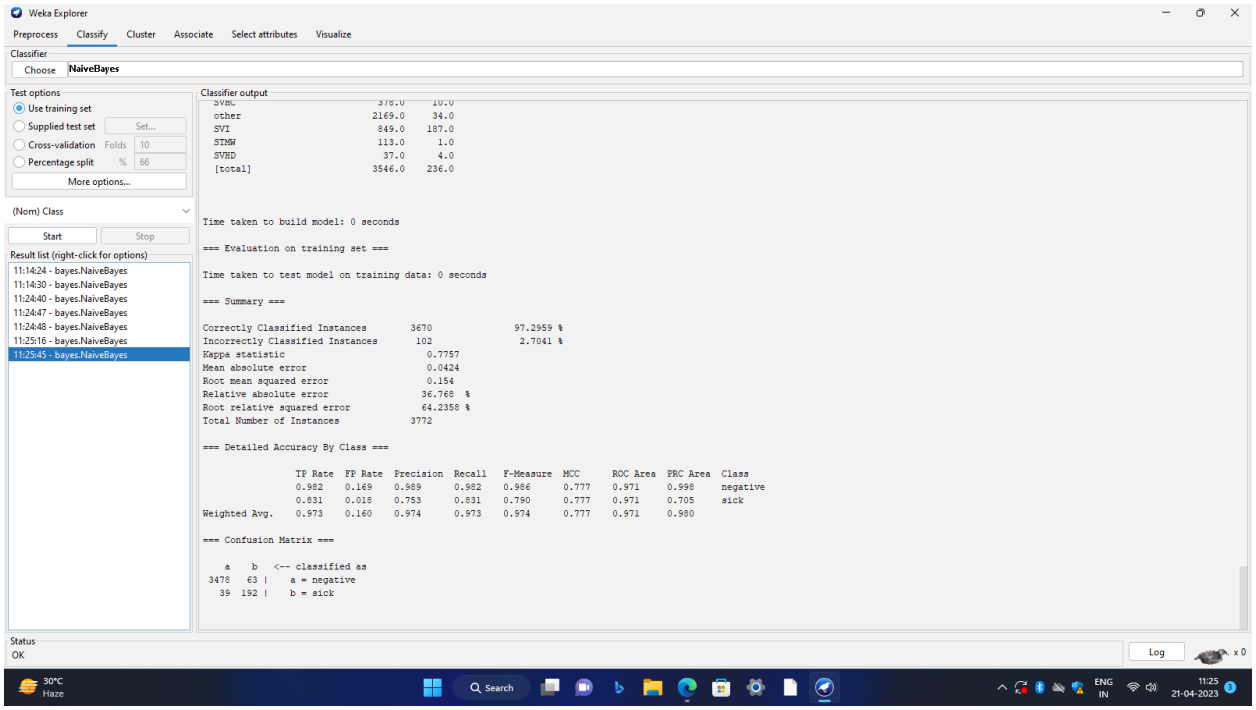
2.7.1


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)
11:14:24 - bayes.NaiveBayes
11:14:30 - bayes.NaiveBayes
11:24:40 - bayes.NaiveBayes
11:24:47 - bayes.NaiveBayes
11:24:48 - bayes.NaiveBayes
Classifier output
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 seconds
=== Evaluation on training set ===
Time taken to test model on training data: 0 seconds
=== Summary ===
Correctly Classified Instances 3461 91.755 %
Incorrectly Classified Instances 311 8.245 %
Kappa statistic 0.3403
Mean absolute error 0.1115
Root mean squared error 0.2395
Relative absolute error 96.844 %
Root relative squared error 99.8797 %
Total Number of Instances 3772
=== Detailed Accuracy By Class ===
TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
0.950 0.580 0.962 0.950 0.956 0.342 0.888 0.992 negative
0.420 0.050 0.354 0.420 0.384 0.342 0.888 0.336 sick
Weighted Avg. 0.918 0.548 0.924 0.918 0.921 0.342 0.888 0.951
=== Confusion Matrix ===
a b <-- classified as
3364 177 | a = negative
134 97 | b = sick
Status
OK
Log
x 0
30°C
Haze
Search
ENG
IN
11:24
21-04-2023

2.7.2


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)
11:14:24 - bayes.NaiveBayes
11:14:30 - bayes.NaiveBayes
11:24:40 - bayes.NaiveBayes
11:24:47 - bayes.NaiveBayes
11:24:48 - bayes.NaiveBayes
11:25:16 - bayes.NaiveBayes
Classifier output
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 seconds
=== Evaluation on training set ===
Time taken to test model on training data: 0.01 seconds
=== Summary ===
Correctly Classified Instances 3666 97.1898 %
Incorrectly Classified Instances 106 2.8102 %
Kappa statistic 0.766
Mean absolute error 0.0456
Root mean squared error 0.1595
Relative absolute error 39.5943 %
Root relative squared error 66.5129 %
Total Number of Instances 3772
=== Detailed Accuracy By Class ===
TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
0.982 0.182 0.988 0.982 0.985 0.767 0.964 0.997 negative
0.818 0.018 0.747 0.818 0.781 0.767 0.964 0.701 sick
Weighted Avg. 0.972 0.172 0.973 0.972 0.972 0.767 0.964 0.979
=== Confusion Matrix ===
a b <-- classified as
3477 64 | a = negative
42 189 | b = sick
Status
OK
Log
x 0
30°C
Haze
Search
ENG
IN
11:25
21-04-2023

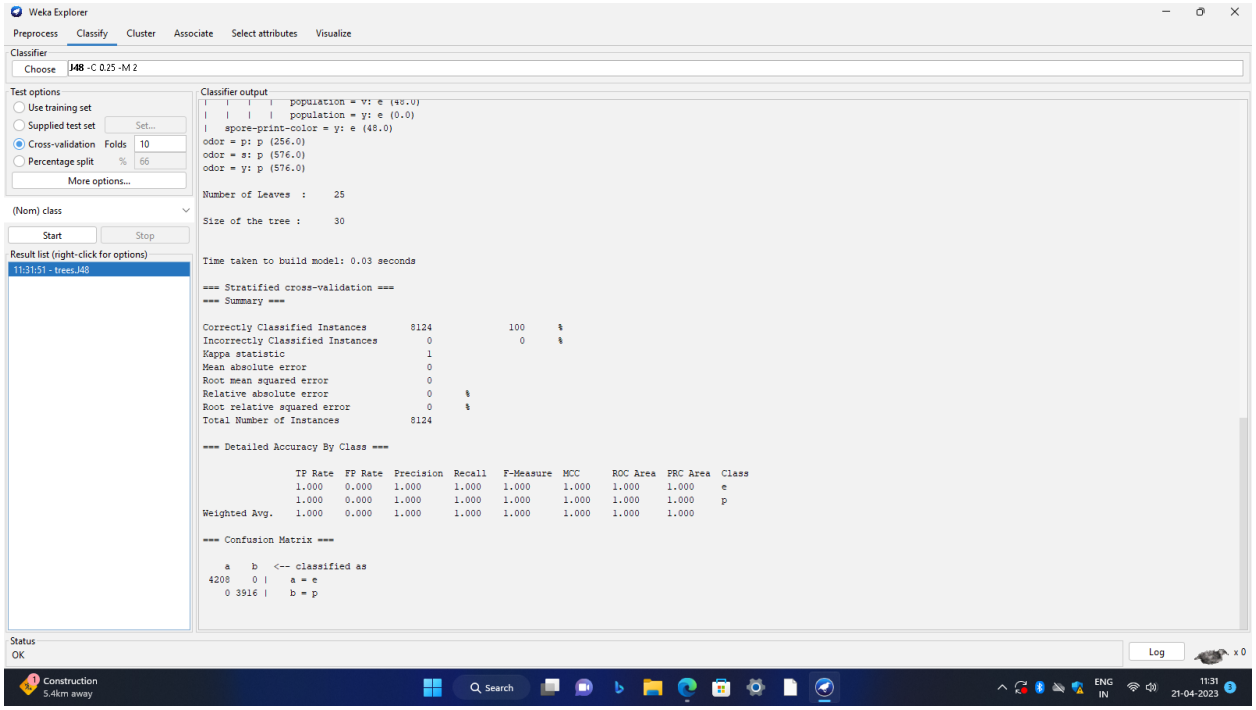
2.7.3

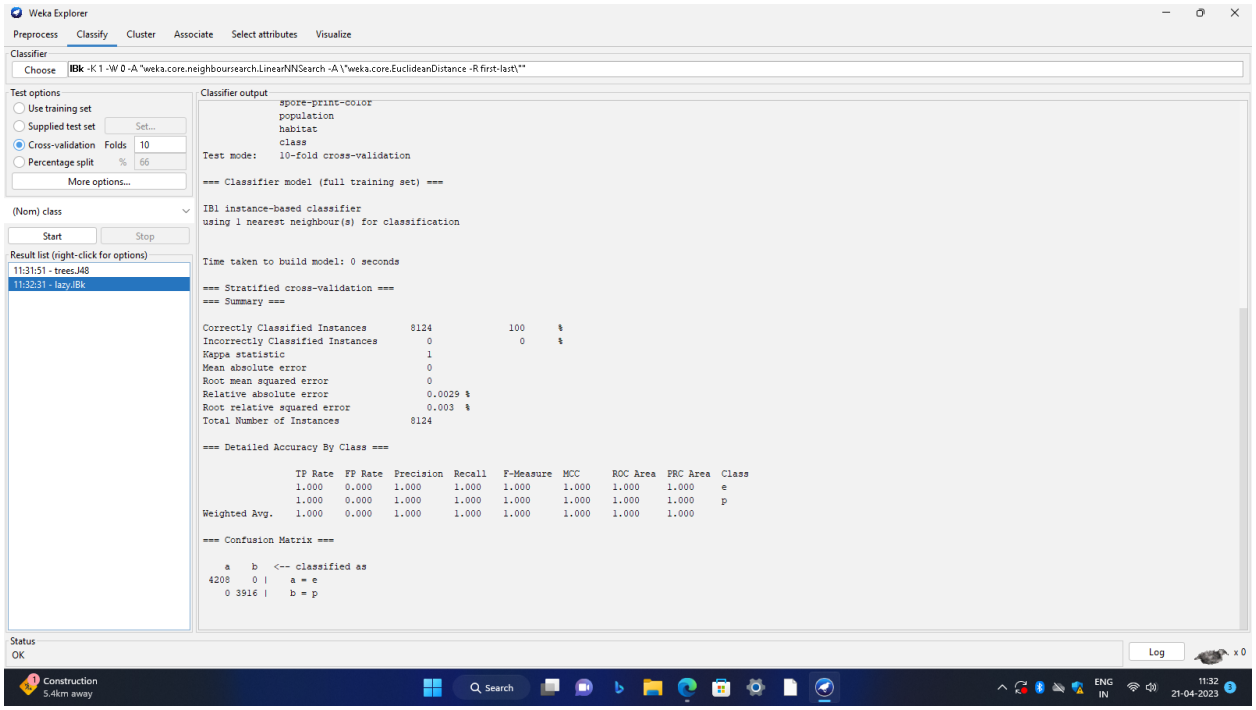


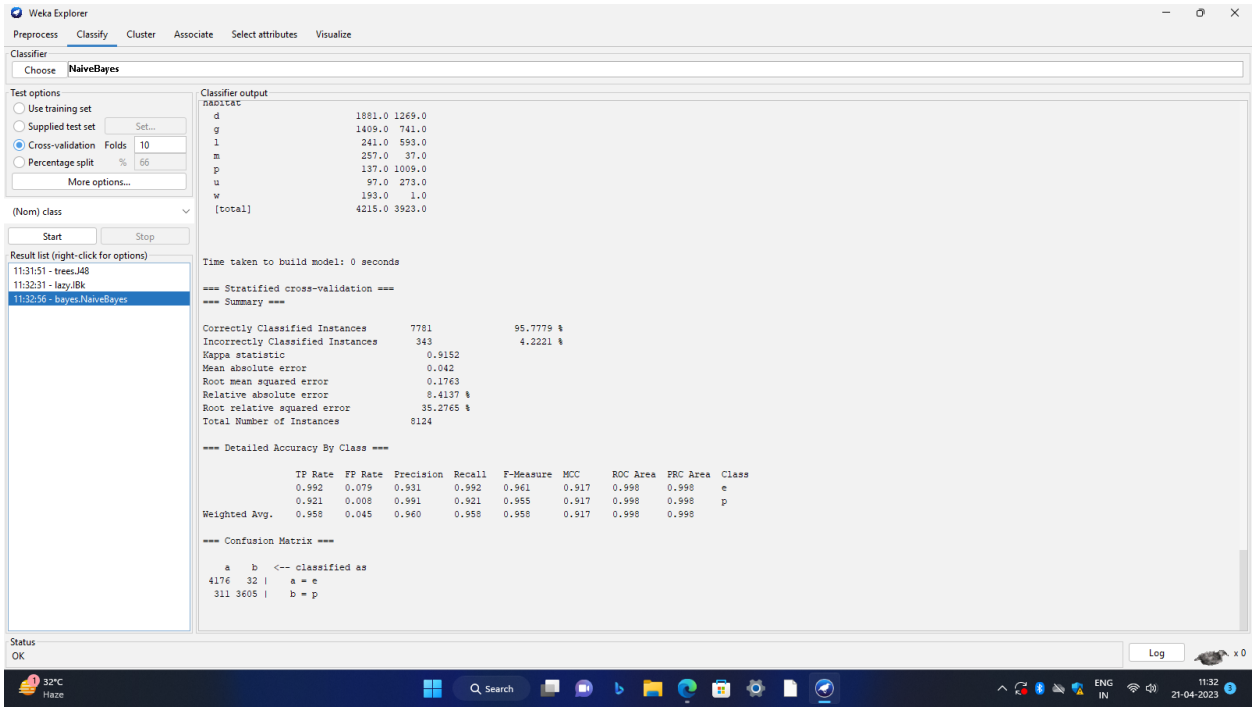
2.8.1

Part II

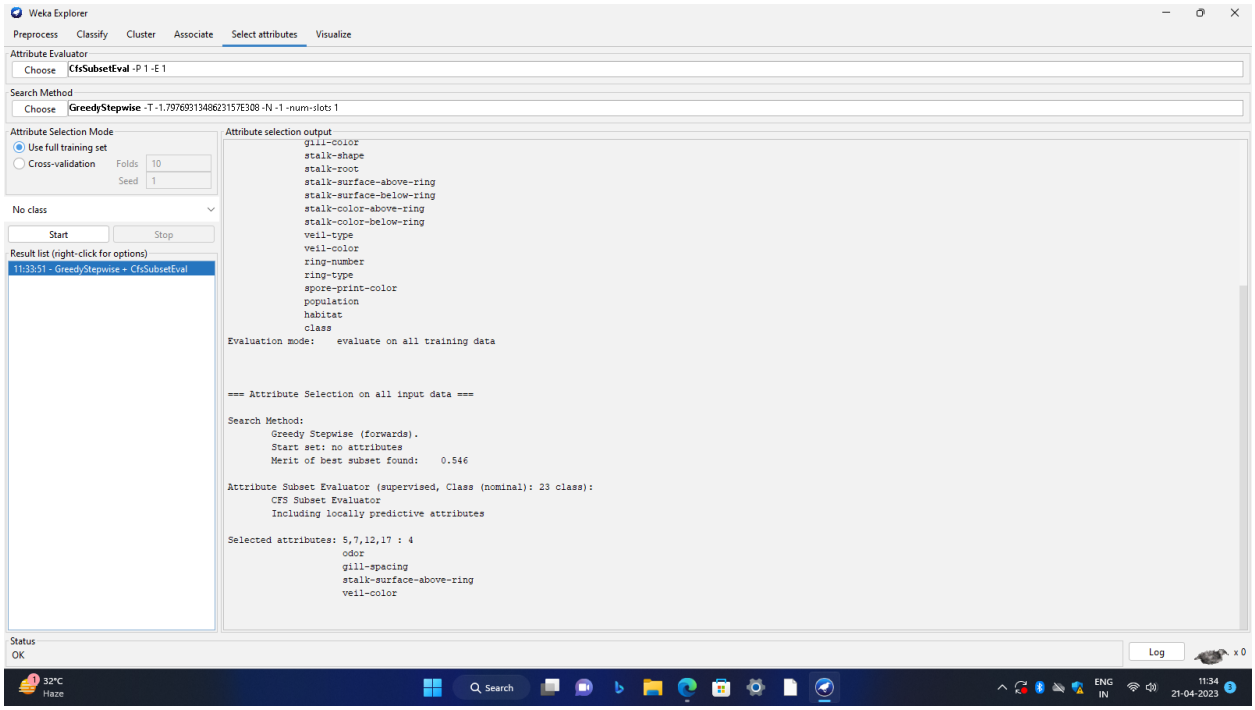
1.2




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)
11:31:51 - trees.J48
11:32:31 - lazy.IBk
Classifier output
spore-print-color
population
habitat
class
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 8124 100 %
Incorrectly Classified Instances 0 0 %
Kappa statistic 1
Mean absolute error 0
Root mean squared error 0
Relative absolute error 0.0029 %
Root relative squared error 0.003 %
Total Number of Instances 8124
=== Detailed Accuracy By Class ===
TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000 e
1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000 p
Weighted Avg. 1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000
=== Confusion Matrix ===
a b <-- classified as
4208 0 | a = e
0 3916 | b = p
Status
OK
Log
x 0
Construction
5.4km away
Search
ENG
IN
11:32
21-04-2023


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)
11:31:51 - trees.J48
11:32:31 - lazy.IBk
11:32:56 - bayes.NaiveBayes
Classifier output
habitat
d 1881.0 1269.0
g 1409.0 741.0
l 241.0 593.0
m 257.0 37.0
p 137.0 1009.0
u 97.0 273.0
w 193.0 1.0
[total] 4215.0 3923.0
Time taken to build model: 0 seconds
=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances 7781 95.7779 %
Incorrectly Classified Instances 343 4.2221 %
Kappa statistic 0.9152
Mean absolute error 0.042
Root mean squared error 0.1763
Relative absolute error 8.4137 %
Root relative squared error 35.2765 %
Total Number of Instances 8124
=== Detailed Accuracy By Class ===
TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
0.992 0.079 0.931 0.992 0.961 0.917 0.998 0.998 e
0.921 0.008 0.991 0.921 0.955 0.917 0.998 0.998 p
Weighted Avg. 0.958 0.045 0.960 0.958 0.958 0.917 0.998 0.998
=== Confusion Matrix ===
a b <-- classified as
4176 32 | a = e
311 3605 | b = p
Status
OK
Log
x 0
32°C
Haze
Search
ENG
IN
11:32
21-04-2023

2.4


Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Attribute Evaluator
Choose
CfsSubsetEval -P 1 -E 1
Search Method
Choose
GreedyStepwise -T -1.7976931348623157E308 -N -1 -num-slots 1
Attribute Selection Mode
Use full training set
Cross-validation
Folds
10
Seed
1
No class
Start
Stop
Result list (right-click for options)
11:33:51 - GreedyStepwise + CfsSubsetEval
Attribute selection output
gill-color
stalk-shape
stalk-root
stalk-surface-above-ring
stalk-surface-below-ring
stalk-color-above-ring
stalk-color-below-ring
veil-type
veil-color
ring-number
ring-type
spore-print-color
population
habitat
class
Evaluation mode: evaluate on all training data
=== Attribute Selection on all input data ===
Search Method:
Greedy Stepwise (forwards).
Start set: no attributes
Merit of best subset found: 0.546
Attribute Subset Evaluator (supervised, Class (nominal): 23 class):
CFS Subset Evaluator
Including locally predictive attributes
Selected attributes: 5,7,12,17 : 4
odor
gill-spacing
stalk-surface-above-ring
veil-color
Status
OK
Log
x 0
32°C
Haze
Search
ENG
IN
11:34
21-04-2023