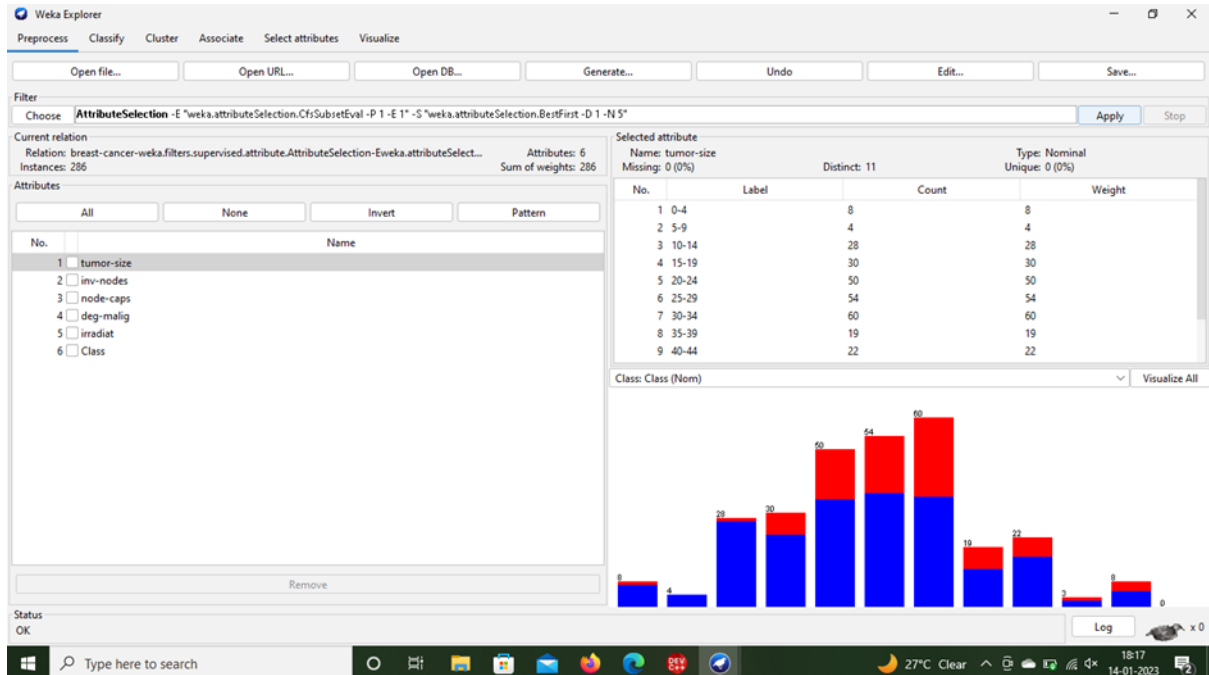


CSA1668-DATAWAREHOUSING AND DATA MINING

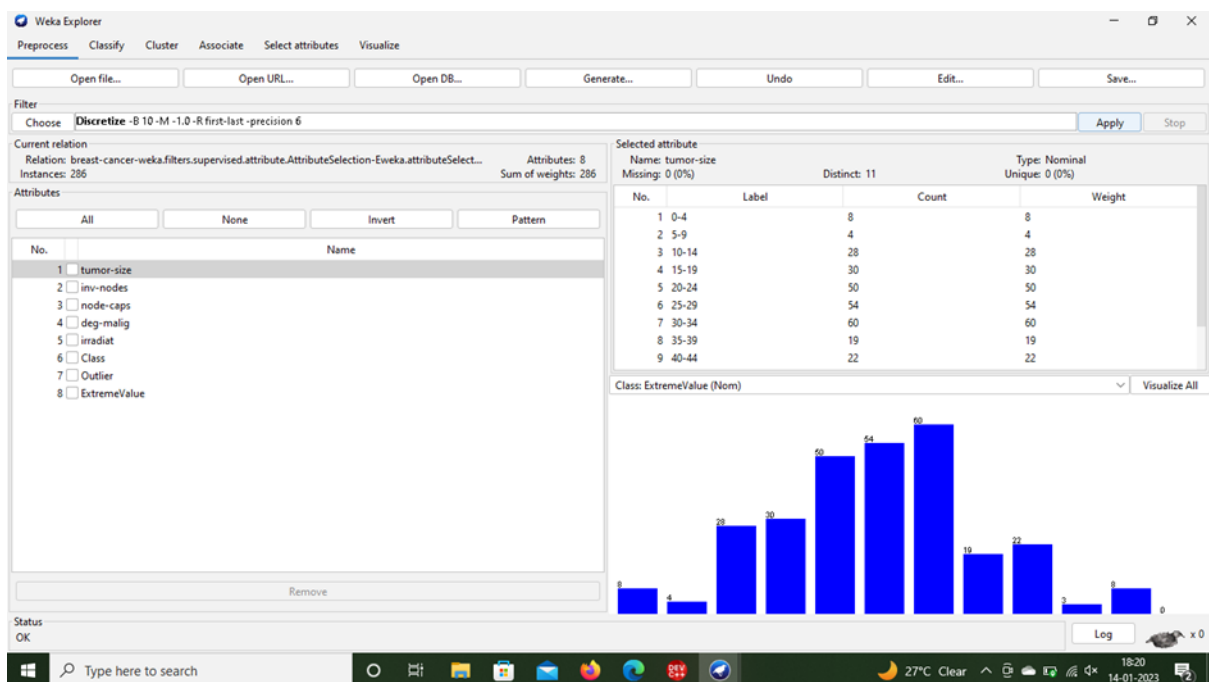
E.SRI HARI

192124168

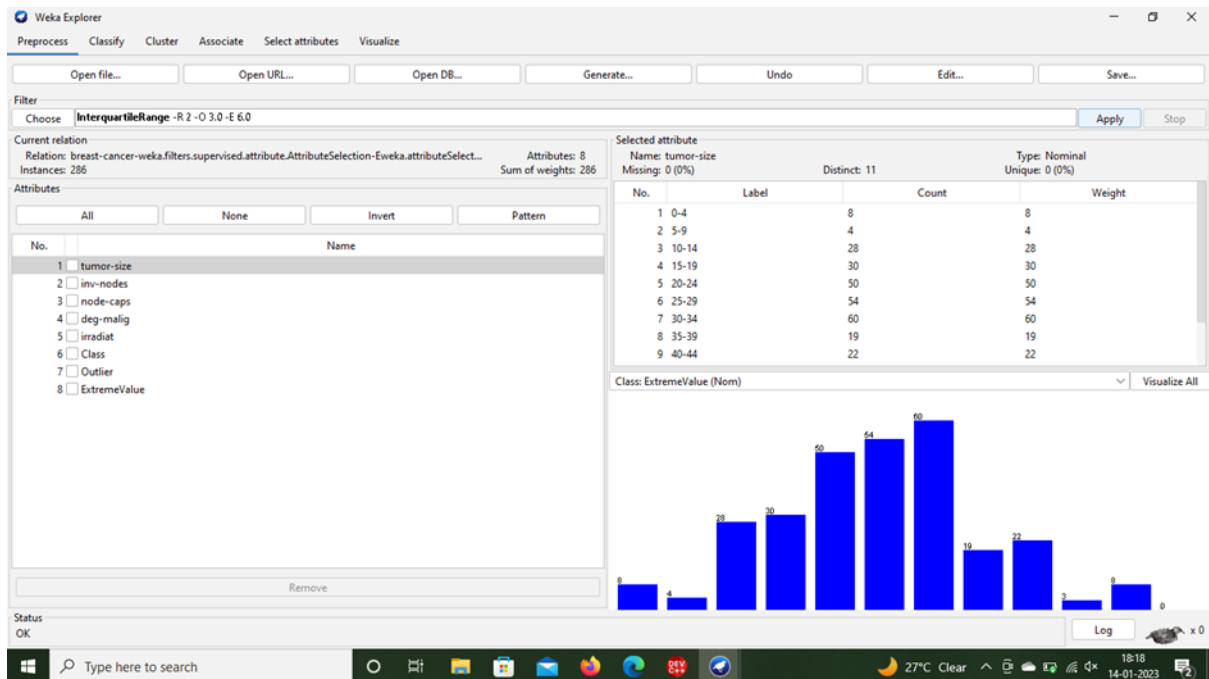
(A)DATA PREPROCESSING



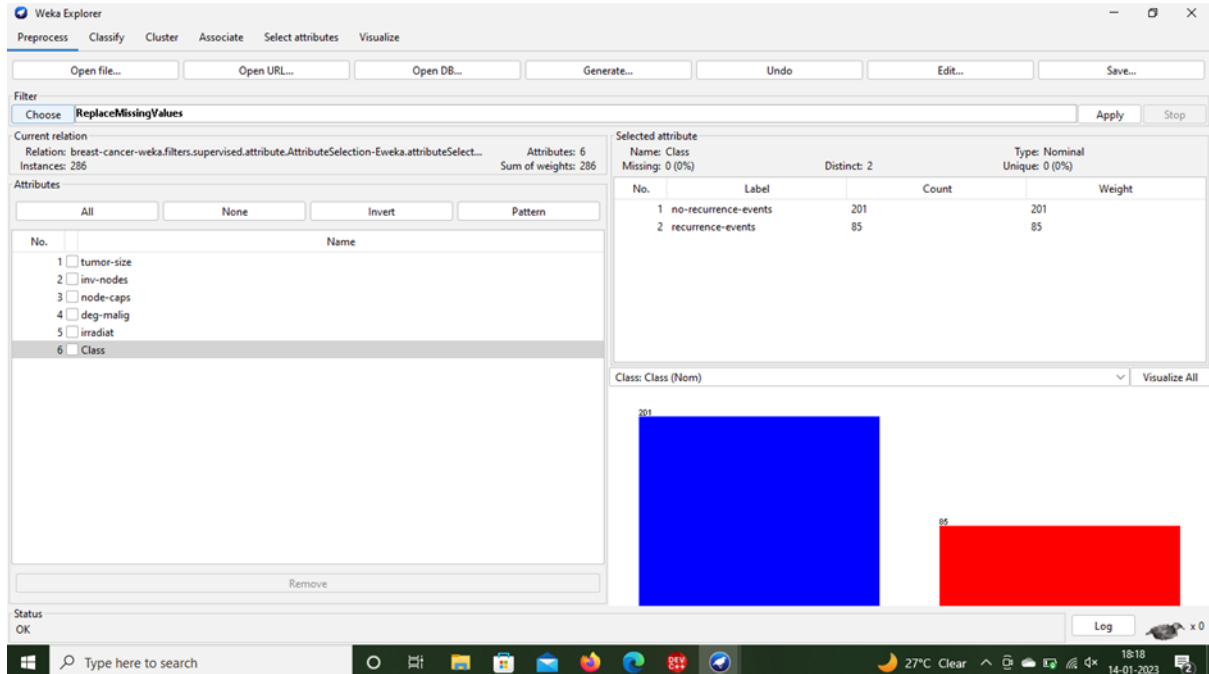
DISCRETIZE:



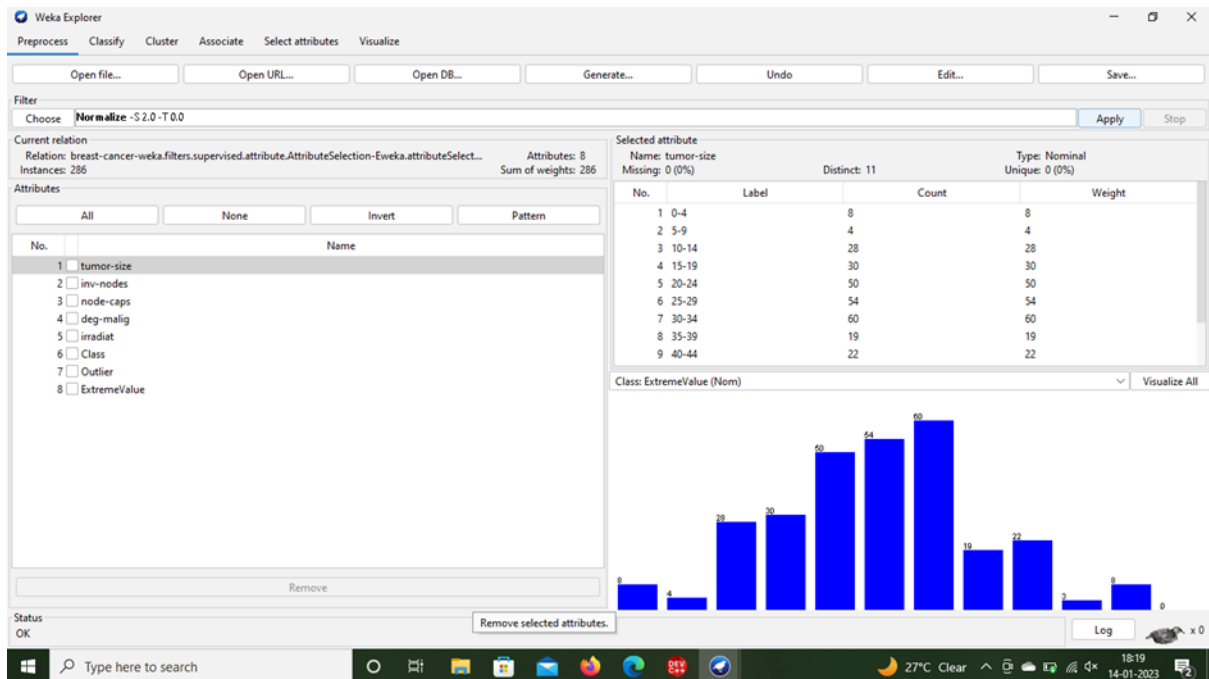
INTERQUARTILE RANGE



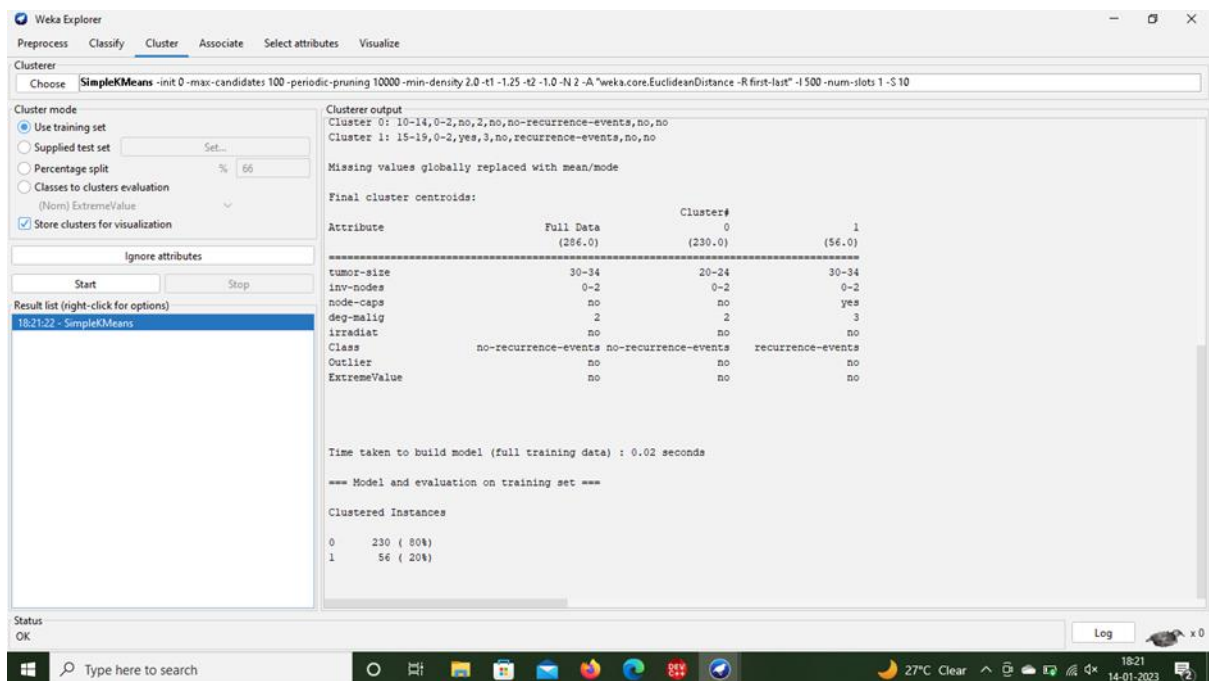
REPLACE MISSING VALUES:



NORMALIZE



CLUSTERING



EM

The screenshot shows the Weka Explorer interface with the 'Cluster' tab selected. The 'Clusterer' dropdown is set to 'EM'. The 'Cluster mode' section has 'Use training set' selected. The 'Ignore attributes' section is empty. The 'Result list' on the left shows '18:21:36 - EM' selected. The 'Clusterer output' pane displays the following information:

Clusterer output

| | | |
|----------------------|----------|---------|
| [total] | 212.8894 | 79.1106 |
| irradiat | | |
| yes | 27.0213 | 42.9787 |
| no | 184.8681 | 35.1319 |
| [total] | 211.8894 | 78.1106 |
| Class | | |
| no-recurrence-events | 169.7021 | 33.2979 |
| recurrence-events | 42.1873 | 44.8127 |
| [total] | 211.8894 | 78.1106 |
| Outlier | | |
| no | 210.8894 | 77.1106 |
| yes | 1 | 1 |
| [total] | 211.8894 | 78.1106 |
| ExtremeValue | | |
| no | 210.8894 | 77.1106 |
| yes | 1 | 1 |
| [total] | 211.8894 | 78.1106 |

Time taken to build model (full training data) : 0.93 seconds

=== Model and evaluation on training set ===

Clustered Instances

| | |
|---|------------|
| 0 | 211 (74%) |
| 1 | 75 (26%) |

Log likelihood: -5.35071

ASSOCIATE:

The screenshot shows the Weka Explorer interface with the 'Associate' tab selected. The 'Associator' dropdown is set to 'Apriori'. The 'Start' button is highlighted. The 'Result list' on the left shows '18:21:55 - Apriori' selected. The 'Associator output' pane displays the following information:

==== Associator model (full training set) ====

Apriori

Minimum support: 0.75 (214 instances)
Minimum metric <confidence>: 0.9
Number of cycles performed: 5

Generated sets of large itemsets:

Size of set of large itemsets L(1): 4
Size of set of large itemsets L(2): 5
Size of set of large itemsets L(3): 2

Best rules found:

1. ExtremeValue=no 286 ==> Outlier=no 286 <conf:(1)> lift:(1) lev:(0) [0] conv:(0)
2. Outlier=no 286 ==> ExtremeValue=no 286 <conf:(1)> lift:(1) lev:(0) [0] conv:(0)
3. node-caps=no 230 ==> Outlier=no 230 <conf:(1)> lift:(1) lev:(0) [0] conv:(0)
4. node-caps=no 230 ==> ExtremeValue=no 230 <conf:(1)> lift:(1) lev:(0) [0] conv:(0)
5. node-caps=no ExtremeValue=no 230 ==> Outlier=no 230 <conf:(1)> lift:(1) lev:(0) [0] conv:(0)
6. node-caps=no Outlier=no 230 ==> ExtremeValue=no 230 <conf:(1)> lift:(1) lev:(0) [0] conv:(0)
7. node-caps=no 230 ==> Outlier=no ExtremeValue=no 230 <conf:(1)> lift:(1) lev:(0) [0] conv:(0)
8. irradiat=no 218 ==> Outlier=no 218 <conf:(1)> lift:(1) lev:(0) [0] conv:(0)
9. irradiat=no 218 ==> ExtremeValue=no 218 <conf:(1)> lift:(1) lev:(0) [0] conv:(0)
10. irradiat=no ExtremeValue=no 218 ==> Outlier=no 218 <conf:(1)> lift:(1) lev:(0) [0] conv:(0)

CLASSIFICATION:

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
☒ Cross-validation Folds: **10**
☐ Percentage split % **66**
More options...

(Nom) ExtremeValue
Start Stop

Result list (right-click for options)
18:17:21 - trees.J48
18:20:14 - trees.J48

Classifier output

Size of the tree : 1

Time taken to build model: 0 seconds

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

| | | | |
|----------------------------------|-----|-----|---|
| Correctly Classified Instances | 286 | 100 | % |
| Incorrectly Classified Instances | 0 | 0 | % |
| Kappa statistic | 1 | | |
| Mean absolute error | 0 | | |
| Root mean squared error | 0 | | |
| Relative absolute error | 0 | % | |
| Root relative squared error | 0 | % | |
| Total Number of Instances | 286 | | |

=== Detailed Accuracy By Class ===

| | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | FRC Area | Class |
|---------------|---------|---------|-----------|--------|-----------|-----|----------|----------|-------|
| | 1.000 | ? | 1.000 | 1.000 | 1.000 | ? | ? | 1.000 | no |
| | ? | 0.000 | ? | ? | ? | ? | ? | ? | yes |
| Weighted Avg. | 1.000 | ? | 1.000 | 1.000 | 1.000 | ? | ? | 1.000 | |

=== Confusion Matrix ===

| | | |
|-----|---|-------------------|
| a | b | <-- classified as |
| 286 | 0 | a = no |
| 0 | 0 | b = yes |

Status
OK

OneDrive
Screenshot saved
The screenshot was added to your OneDrive.

27°C Clear 18:20 14-01-2023