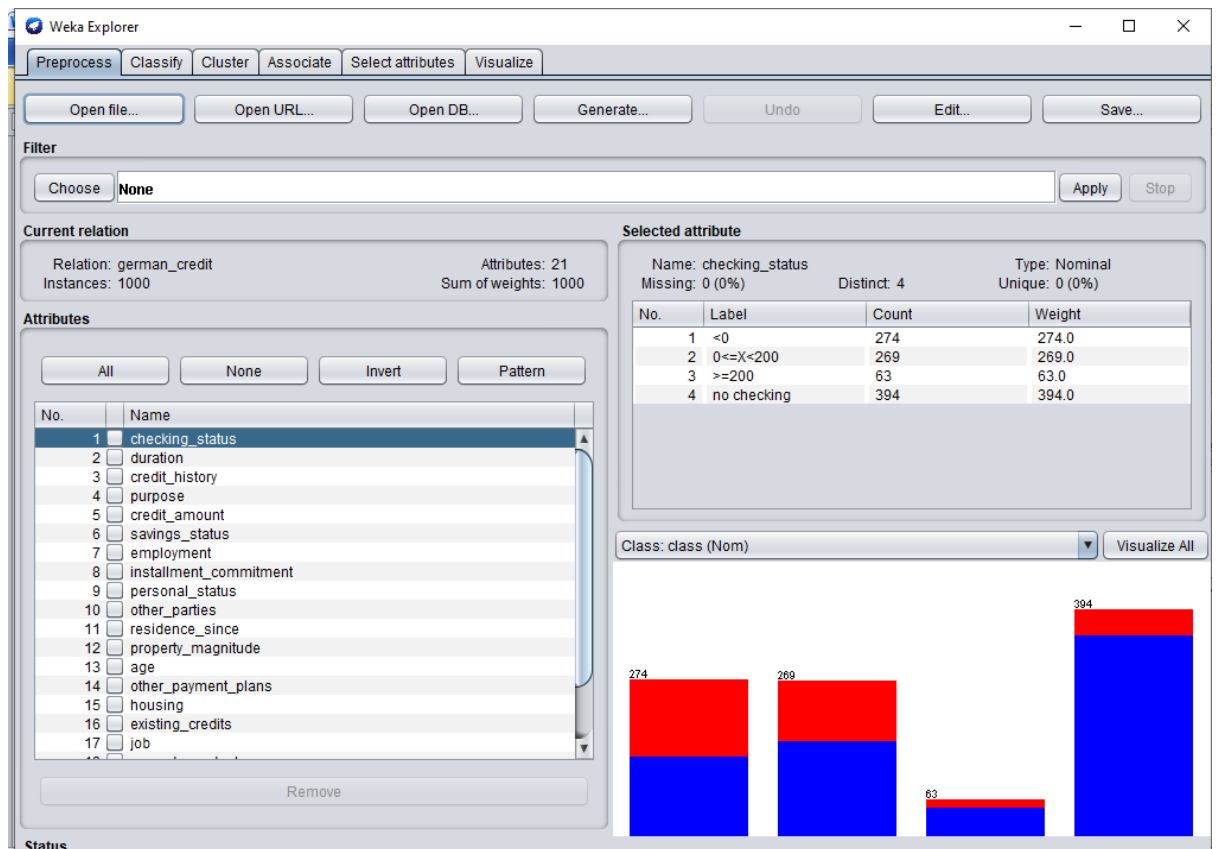


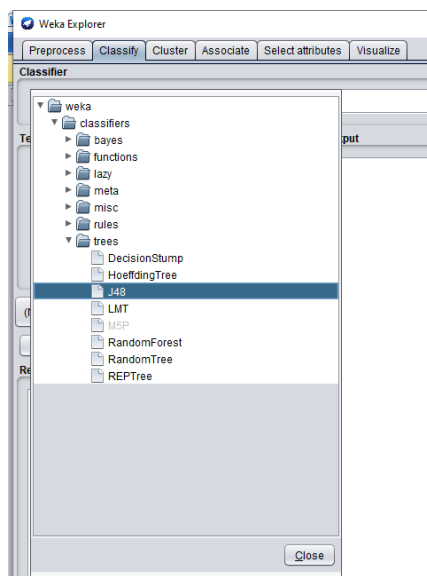
# ASSIGNMENT- 3

## Deriving a Decision Tree from Given Dataset:

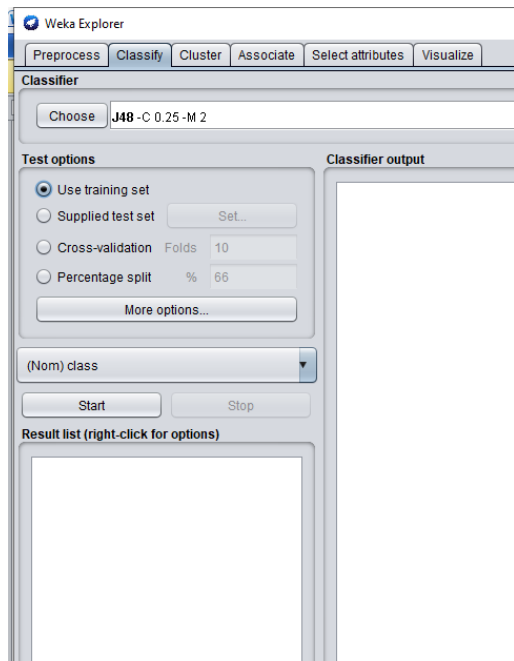
- Open the German Credit Dataset to perform the Decision Tree operation.



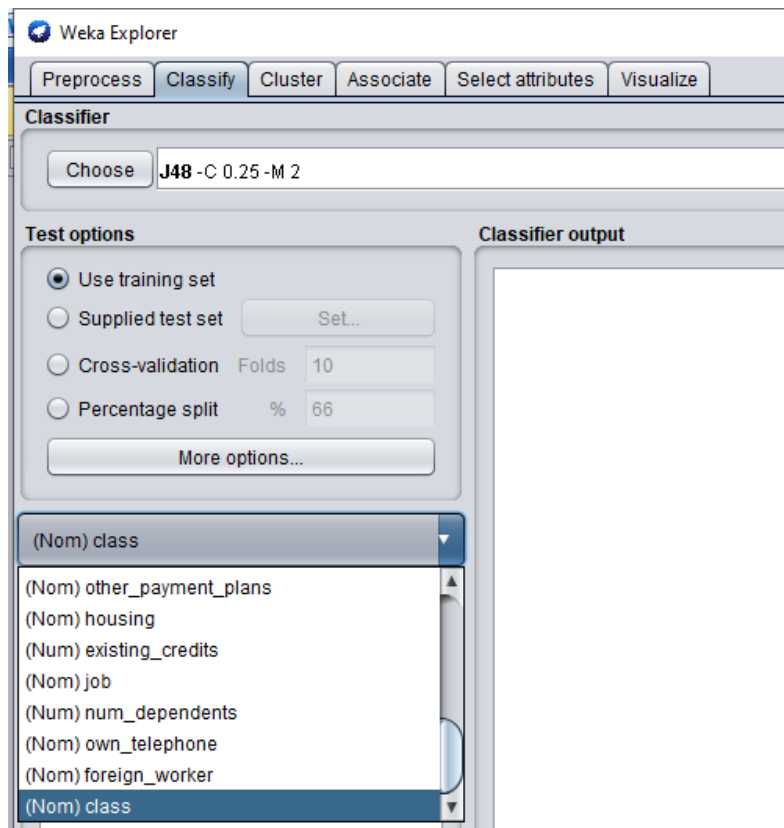
- Click on classify tab and Choose J48 from trees



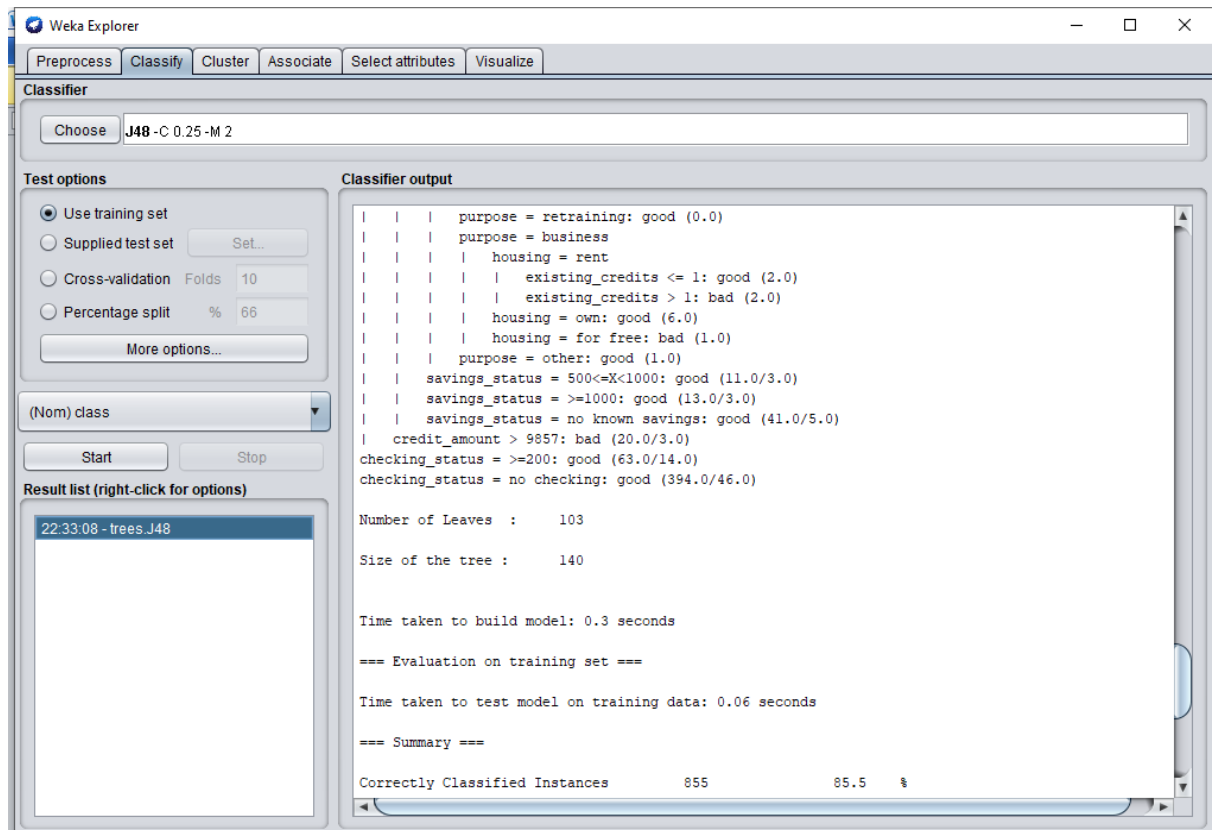
- Choose the Use training set from Test options



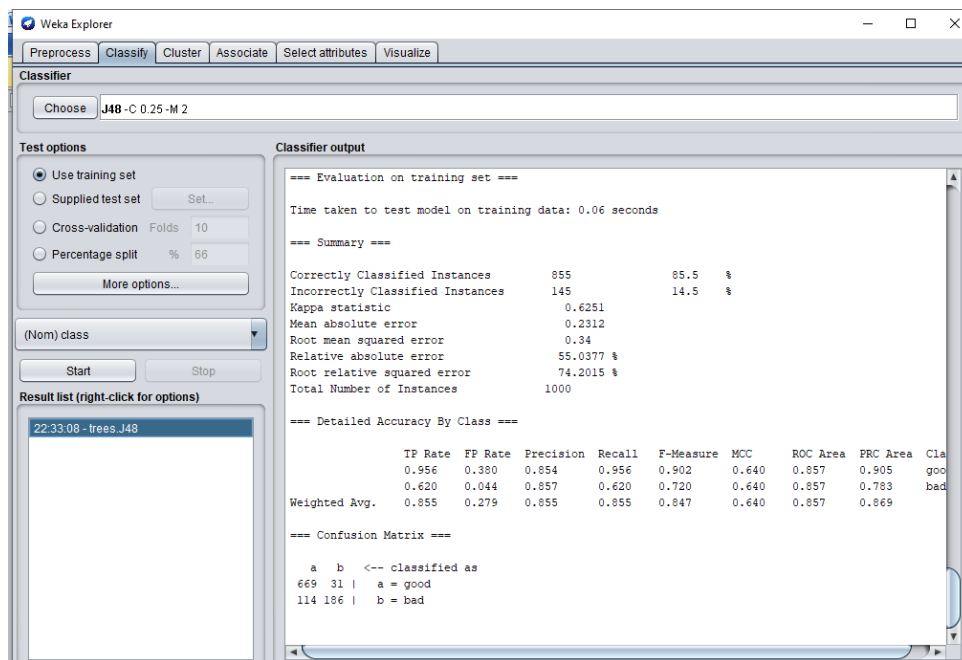
- Now, choose the attribute **class** from the dropdown. By default it is class if not change it to class.



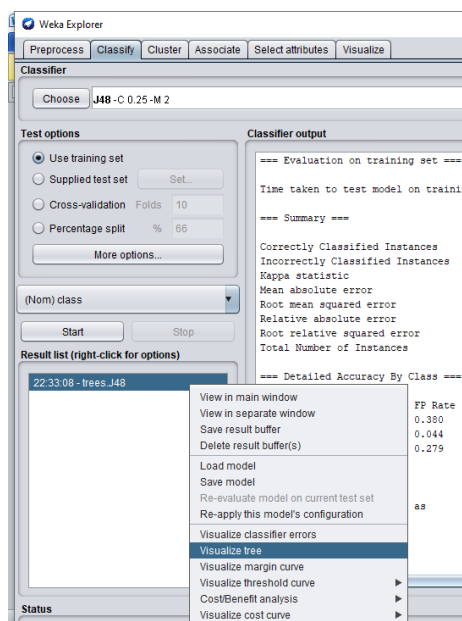
- Next click on Start.
- The output is shown as below.



- Number of leaves = 103
- Number of tree = 140
- Time Taken to build Model = 0.3 sec.
- The Confusion Matrix is show in the below screenshot:

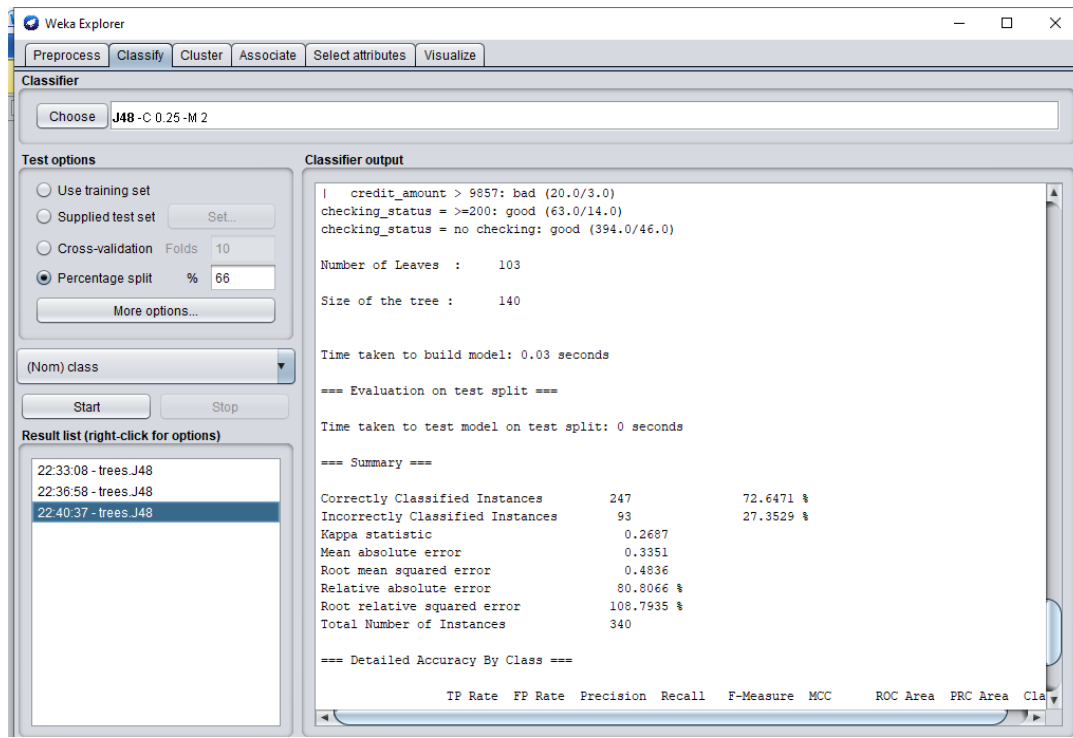


- select the **visualise tree** option by right click on the result list.

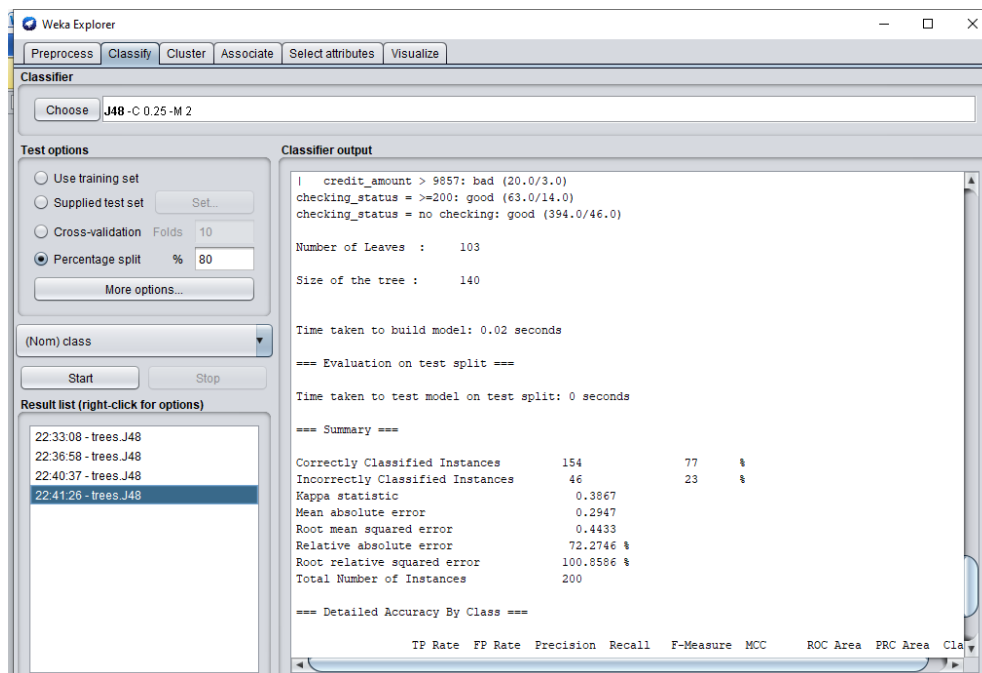


- The Output is shown below:





- Accuracy when Percentage split is 80%



## 2) FOR NAIVEBAYES CLASSIFIER:

- Accuracy when Percentage split is 66%

The screenshot shows the Weka Explorer interface with the NaiveBayes classifier selected. The 'Test options' section has 'Percentage split' set to 66%. The 'Classifier output' pane displays the following results:

```

=== Evaluation on test split ===

Time taken to test model on test split: 0.01 seconds

=== Summary ===

Correctly Classified Instances      260           76.4706 %
Incorrectly Classified Instances    80           23.5294 %
Kappa statistic                    0.3824
Mean absolute error                 0.2819
Root mean squared error             0.4005
Relative absolute error             67.9798 %
Root relative squared error         90.114 %
Total Number of Instances          340

=== Detailed Accuracy By Class ===

          TP Rate  FP Rate  Precision  Recall  F-Measure  MCC   ROC Area  PRC Area  Cla
          0.852    0.478    0.832    0.852    0.842    0.383  0.804    0.921    goo
          0.522    0.148    0.560    0.522    0.540    0.383  0.804    0.592    bad
Weighted Avg.    0.765    0.390    0.760    0.765    0.762    0.383  0.804    0.834

=== Confusion Matrix ===

  a  b  <-- classified as
213  37 |  a = good
 43  47 |  b = bad
  
```

The 'Result list' on the left shows several entries, with '23:32:45 - bayes.NaiveBayes' selected.

- Accuracy when Percentage split is 80%

The screenshot shows the Weka Explorer interface with the NaiveBayes classifier selected. The 'Test options' section has 'Percentage split' set to 80%. The 'Classifier output' pane displays the following results:

```

=== Evaluation on test split ===

Time taken to test model on test split: 0.01 seconds

=== Summary ===

Correctly Classified Instances      149           74.5 %
Incorrectly Classified Instances    51           25.5 %
Kappa statistic                    0.3657
Mean absolute error                 0.2879
Root mean squared error             0.4129
Relative absolute error             70.6169 %
Root relative squared error         93.9316 %
Total Number of Instances          200

=== Detailed Accuracy By Class ===

          TP Rate  FP Rate  Precision  Recall  F-Measure  MCC   ROC Area  PRC Area  Cla
          0.799    0.412    0.850    0.799    0.824    0.368  0.796    0.923    goo
          0.588    0.201    0.500    0.588    0.541    0.368  0.796    0.539    bad
Weighted Avg.    0.745    0.358    0.761    0.745    0.751    0.368  0.796    0.825

=== Confusion Matrix ===

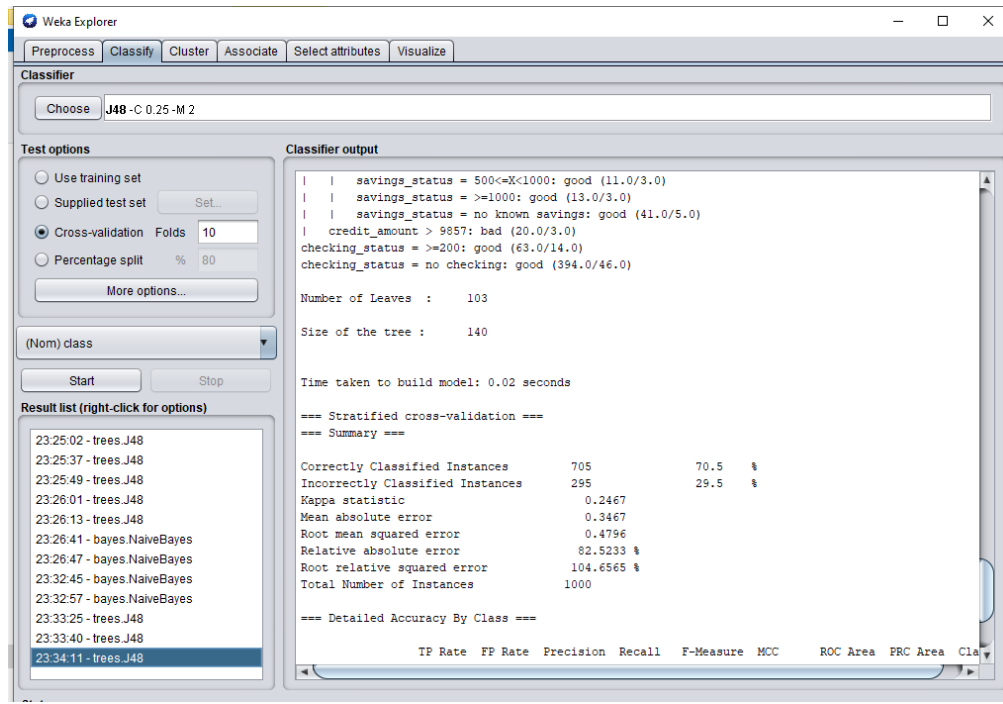
  a  b  <-- classified as
119  30 |  a = good
 21  30 |  b = bad
  
```

The 'Result list' on the left shows several entries, with '23:32:57 - bayes.NaiveBayes' selected.

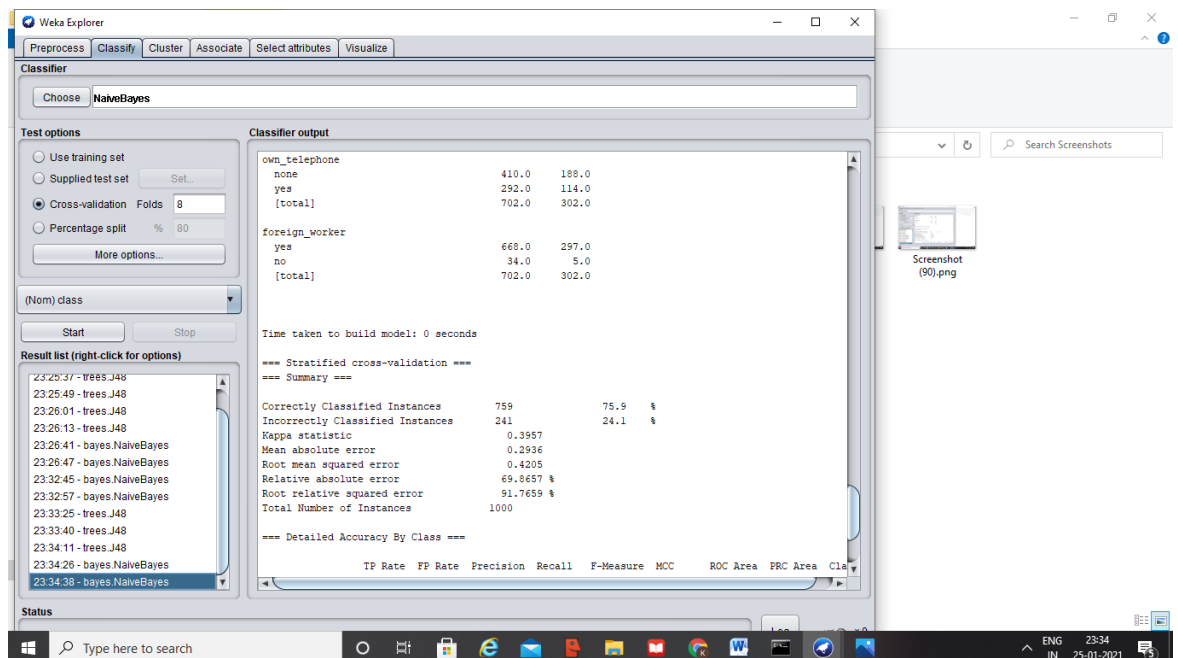
## VARYING THE CROSS VALIDATIONS FOR ACCURACY:

### 1) FOR J48 CLASSIFIER:

- Accuracy when Cross – validation is 10



- Accuracy when Cross – validation is 8



## 2) FOR NAIVEBAYES CLASSIFIER:

- Accuracy when Cross validation is 10





