## DATA MINING ASSIGNMENT 2

# NaiveBayes Classification

**TASK 1**: One type of model that you can create is Naivebayes. Train a Naivebayes using the complete dataset as the training data. Report the model obtained after training.

PROCEDURE:

1) Open Weka GUI Chooser.

2) Select WORKBENCH present in Applications.

3) Go to OPEN file and browse the file that is already stored in the system “credit-g.arff”.

4) Go to the Classify tab.

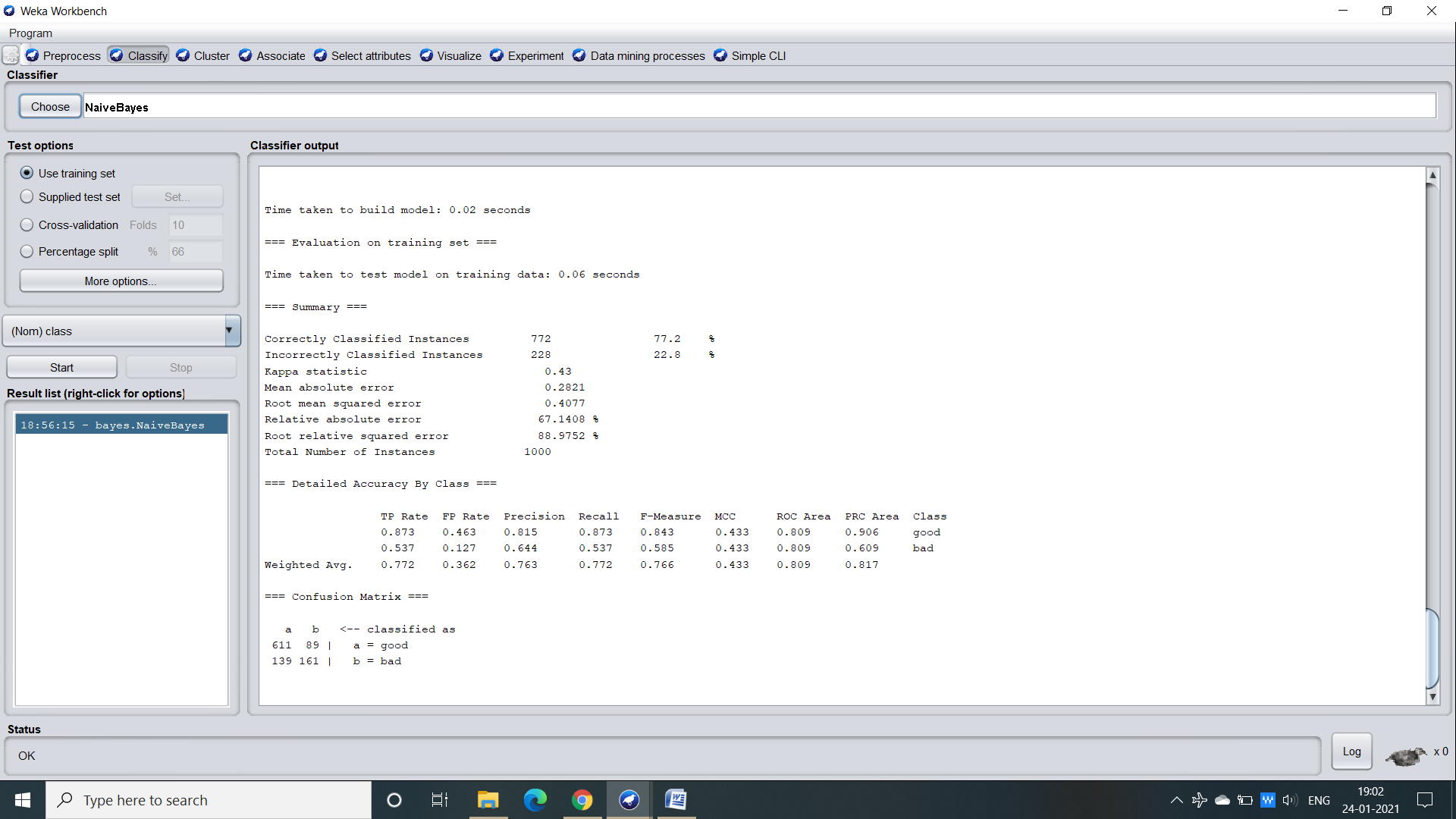
5) Click on the choose button then select NaiveBayes in Bayes dropdown list.

6) Select Test options “Use training set”.

7) Select class attribute.

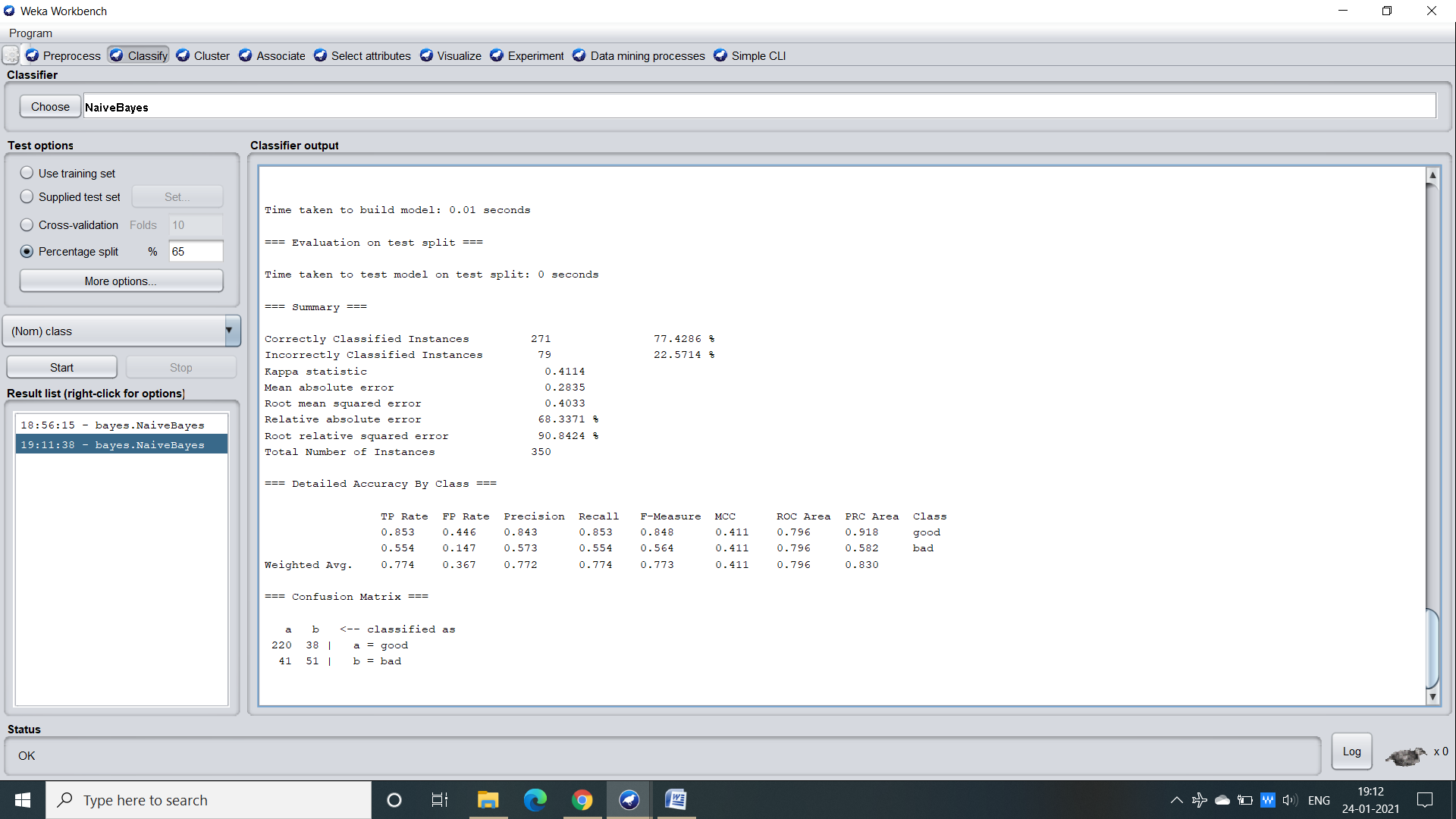
8) Click Start.

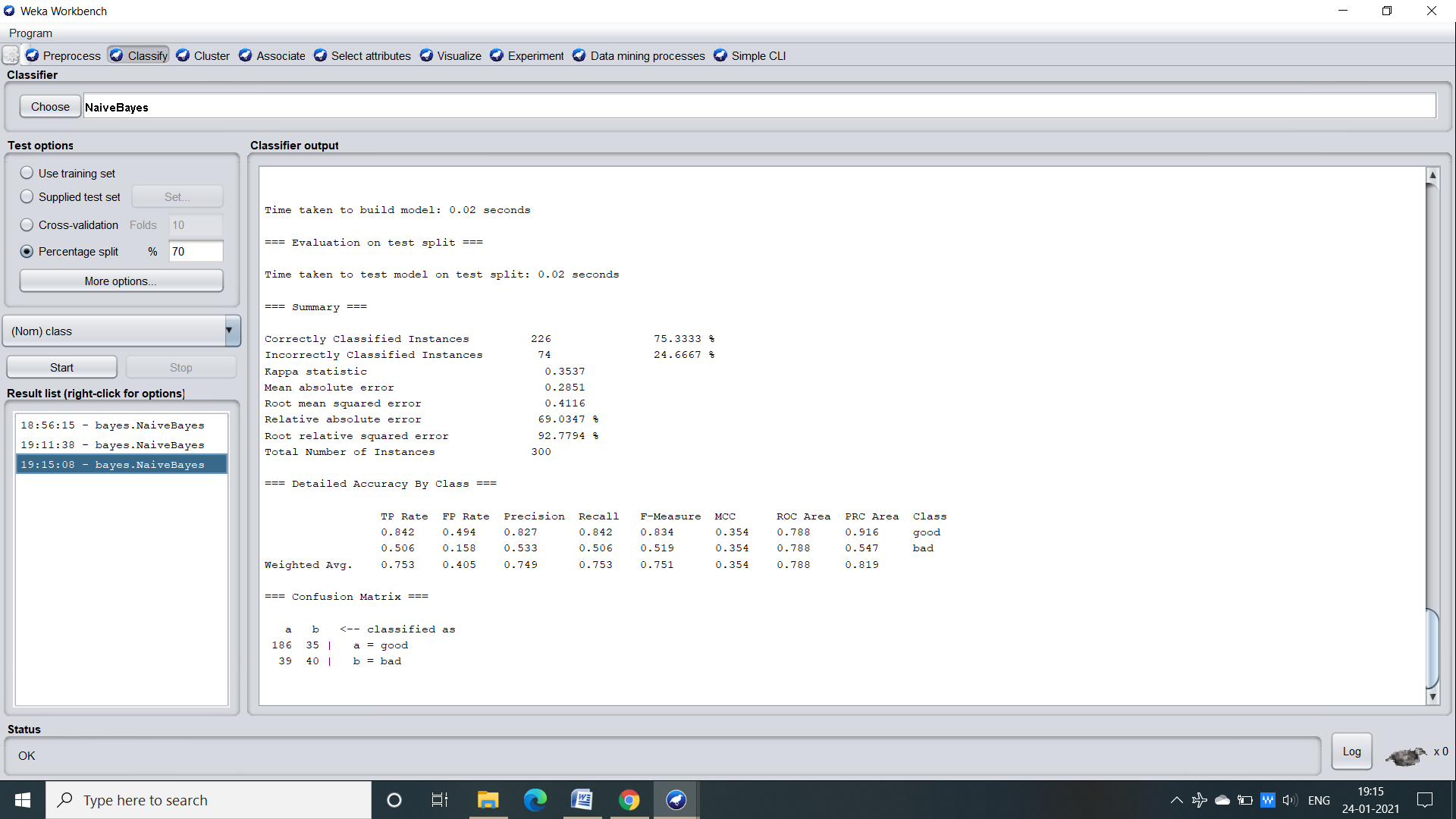
9) Now we can see the output details in the Classifier output.



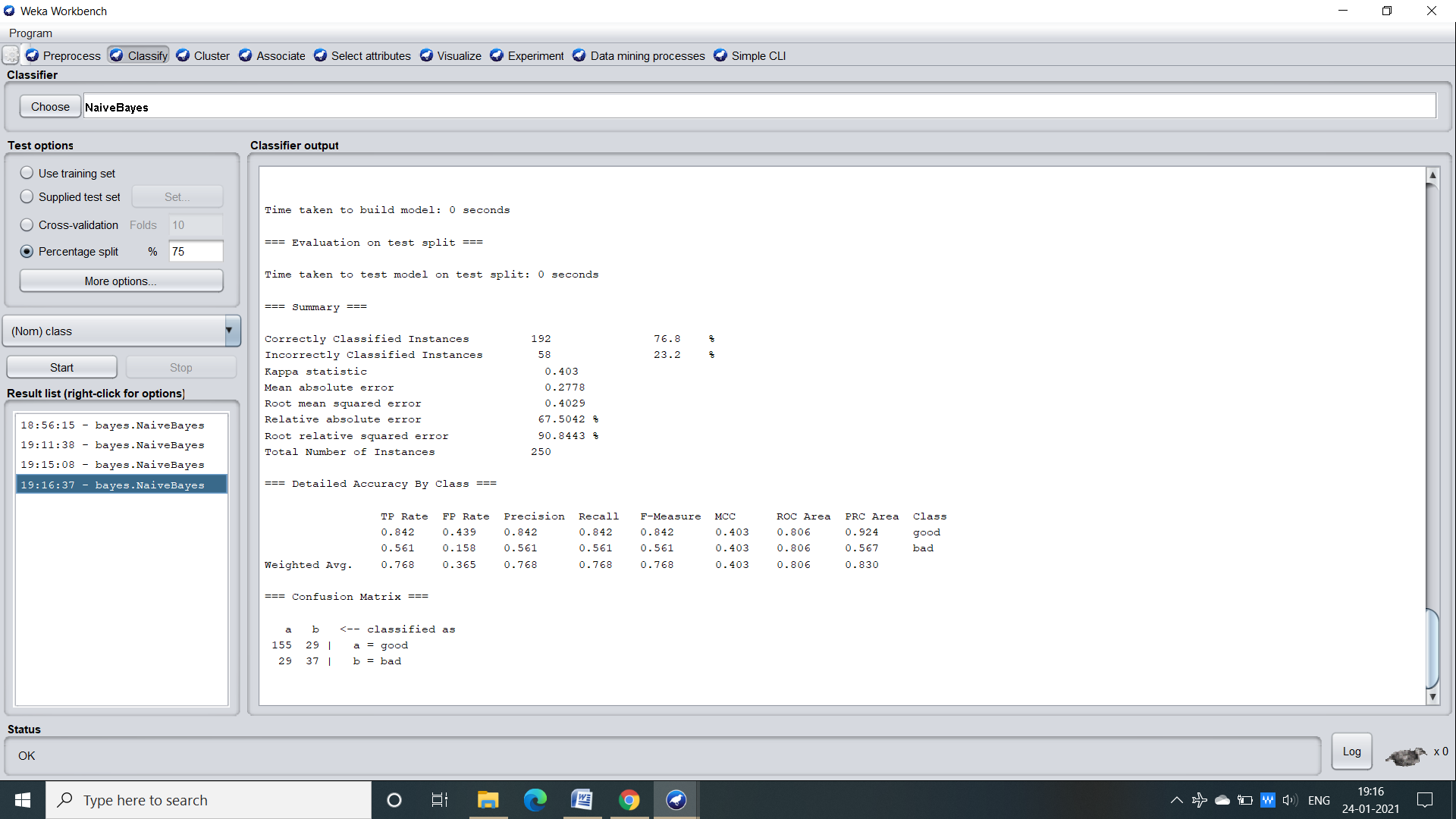
**TASK 2:** Train a NaiveBayes using percentage split and report your results.Increase percentage split by 5% upto 80% starting from 65% and check at which percentage split we are getting the best accuracy.

1.When the percentage split is 65%,the accuracy is 77.4286%.

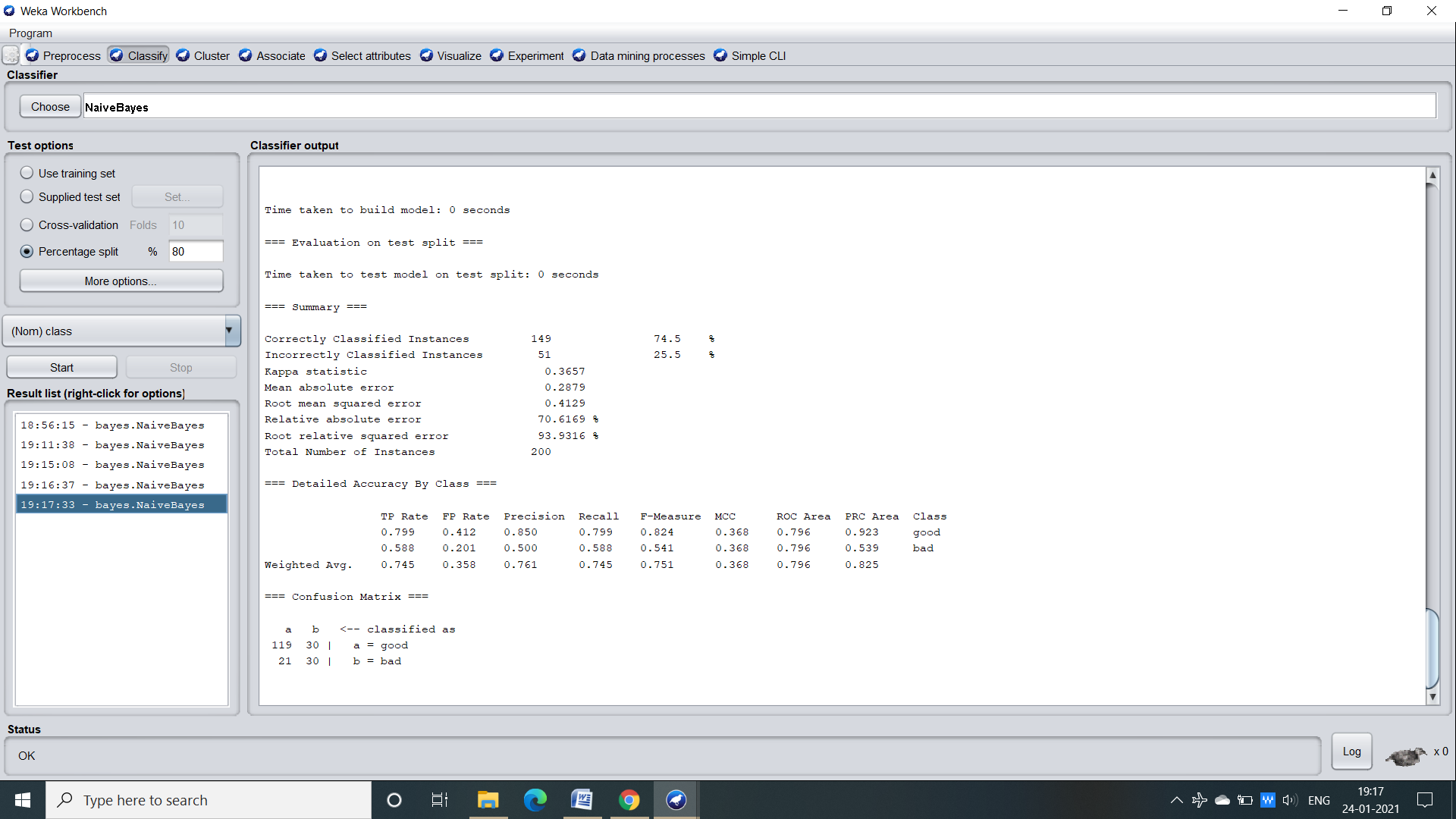


2.When the percentage is 70% the accuracy is 75.3333% 

3.When the percentage is 75% the accuracy is 76.8%



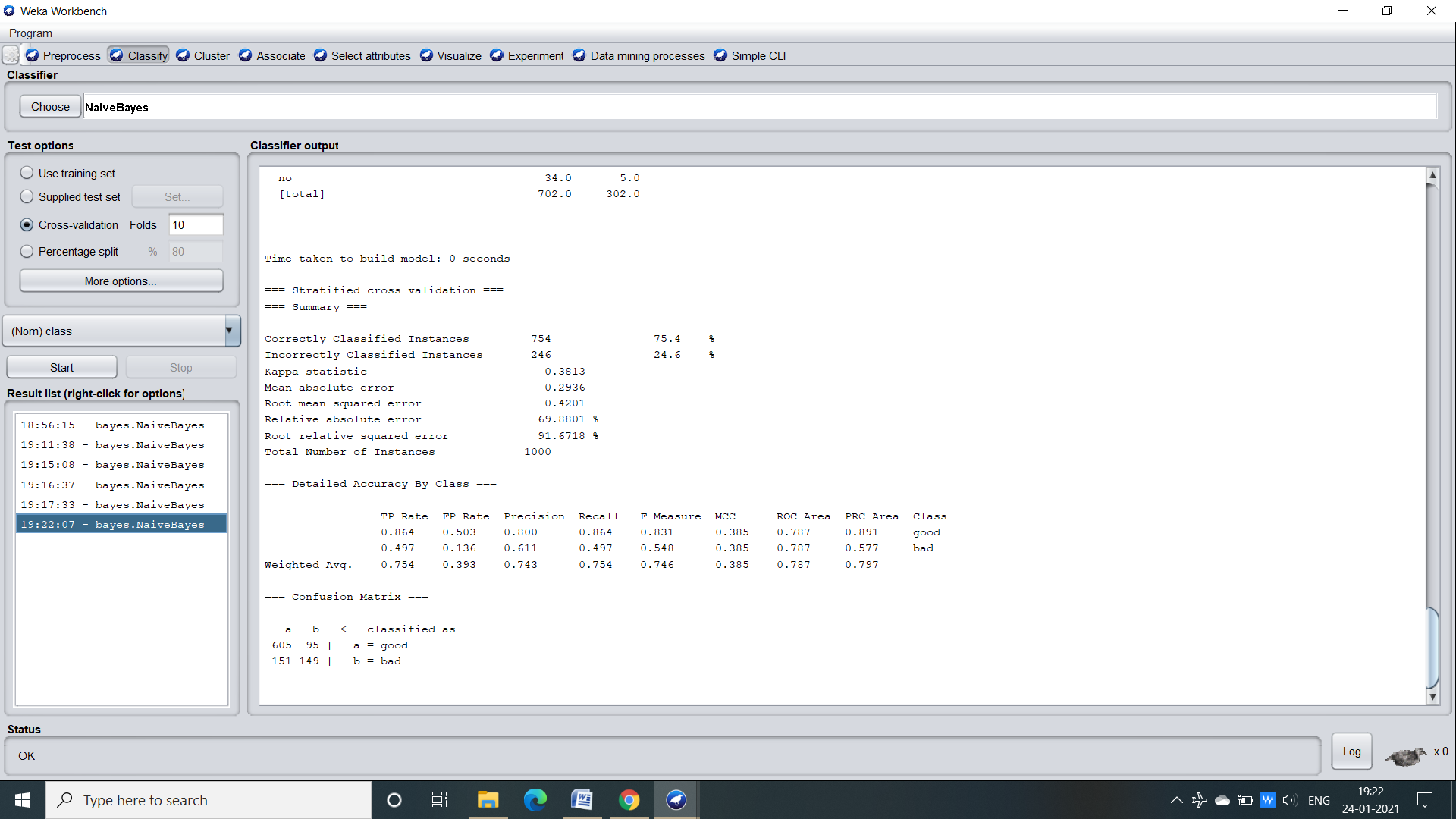
4.When the percentage is 80 the accuracy is 74.5%



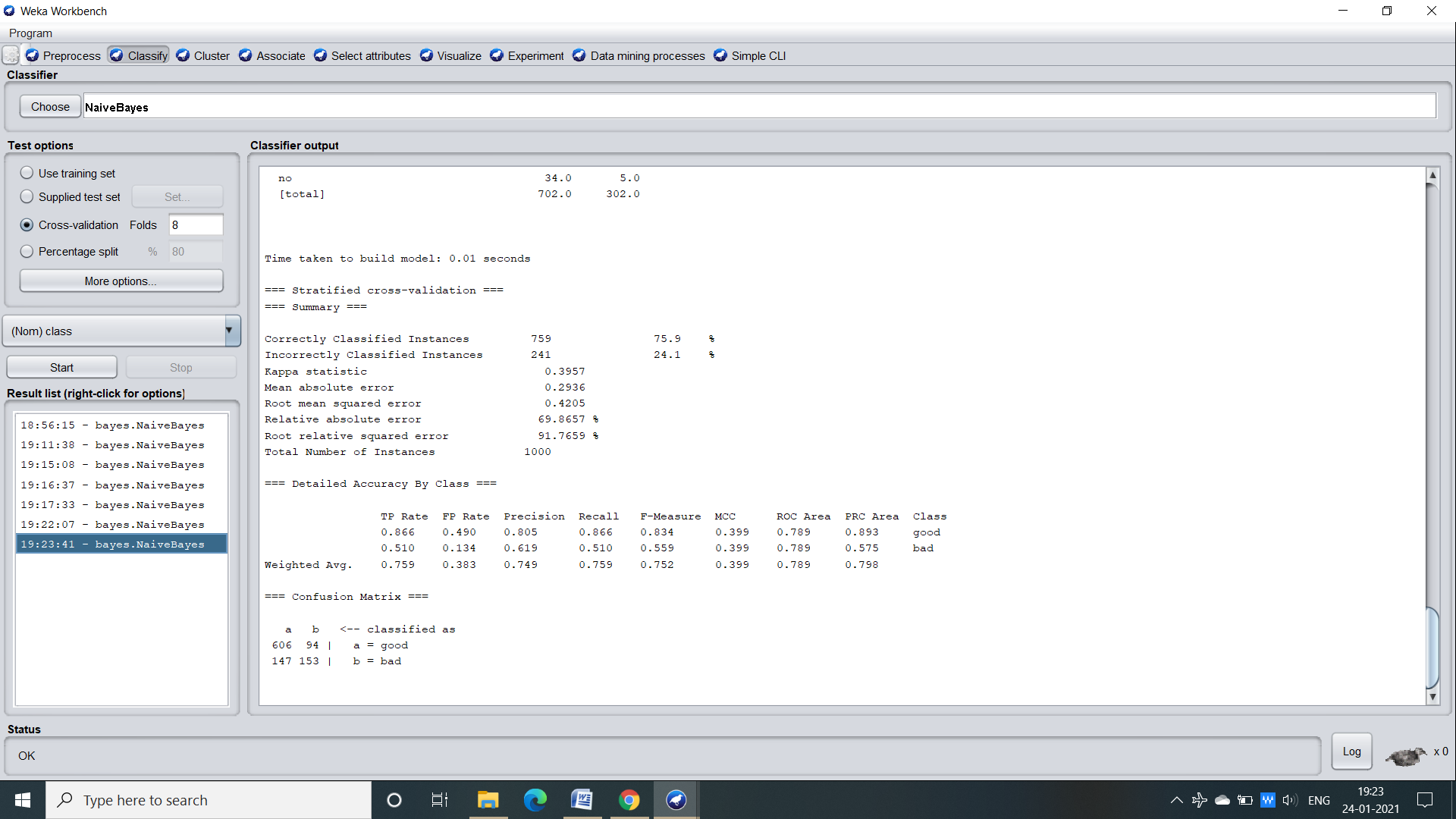
CONCLUSION : When the percentage split is 65%,the accuracy is high which is 77%.

**TASK 3**:Train the NaiveBayes using cross validation and report the results.

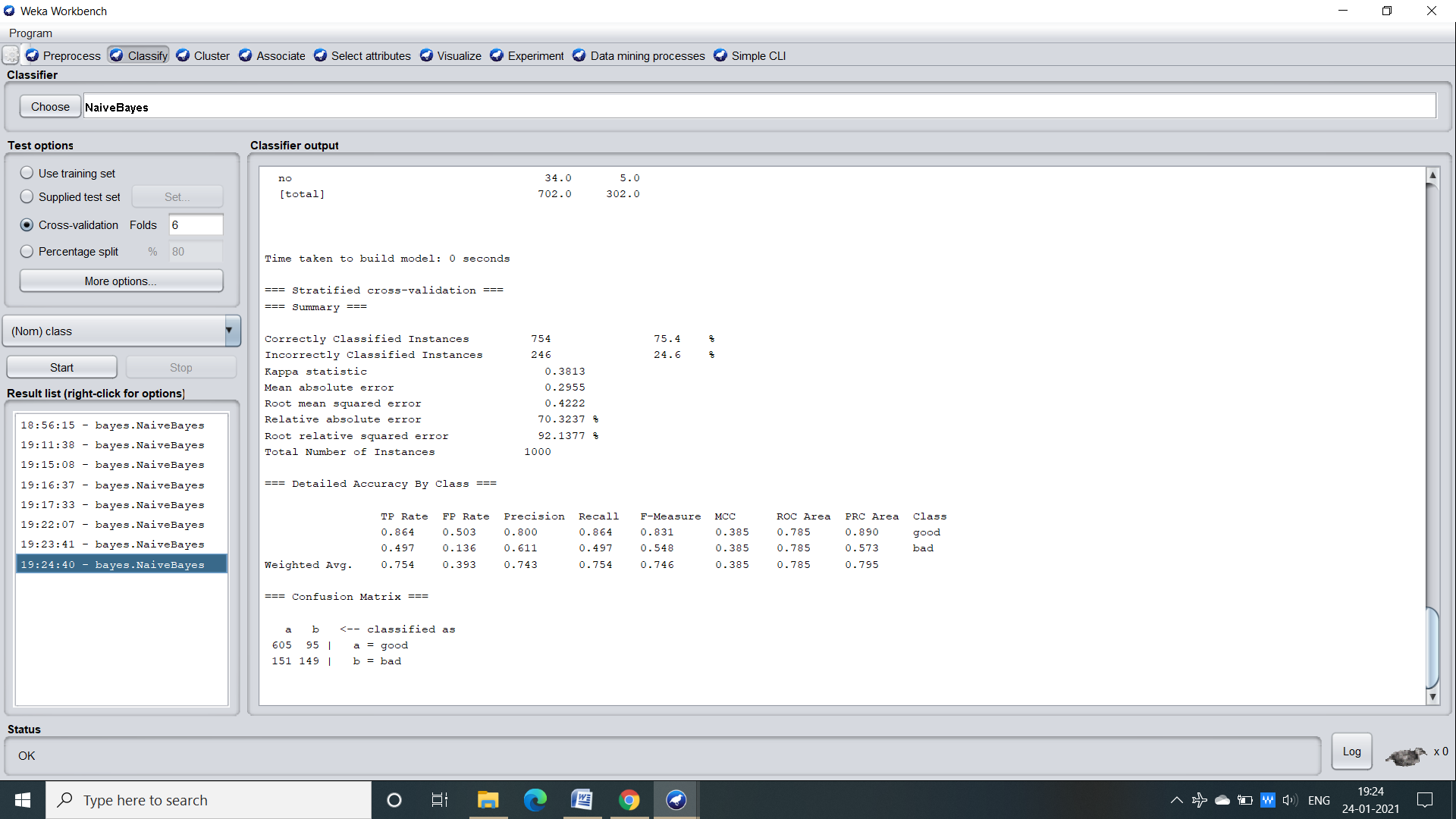
1.When cross validation folds:10,accuracy is 75.4%.



2.When cross validation folds : 8,accuracy is 75.9%.



3.When cross Validation folds:6, accuracy is 75.4%.



CONCLUSION : The accuracy is high when the number of cross validation folds are 8 which is 75.9%.