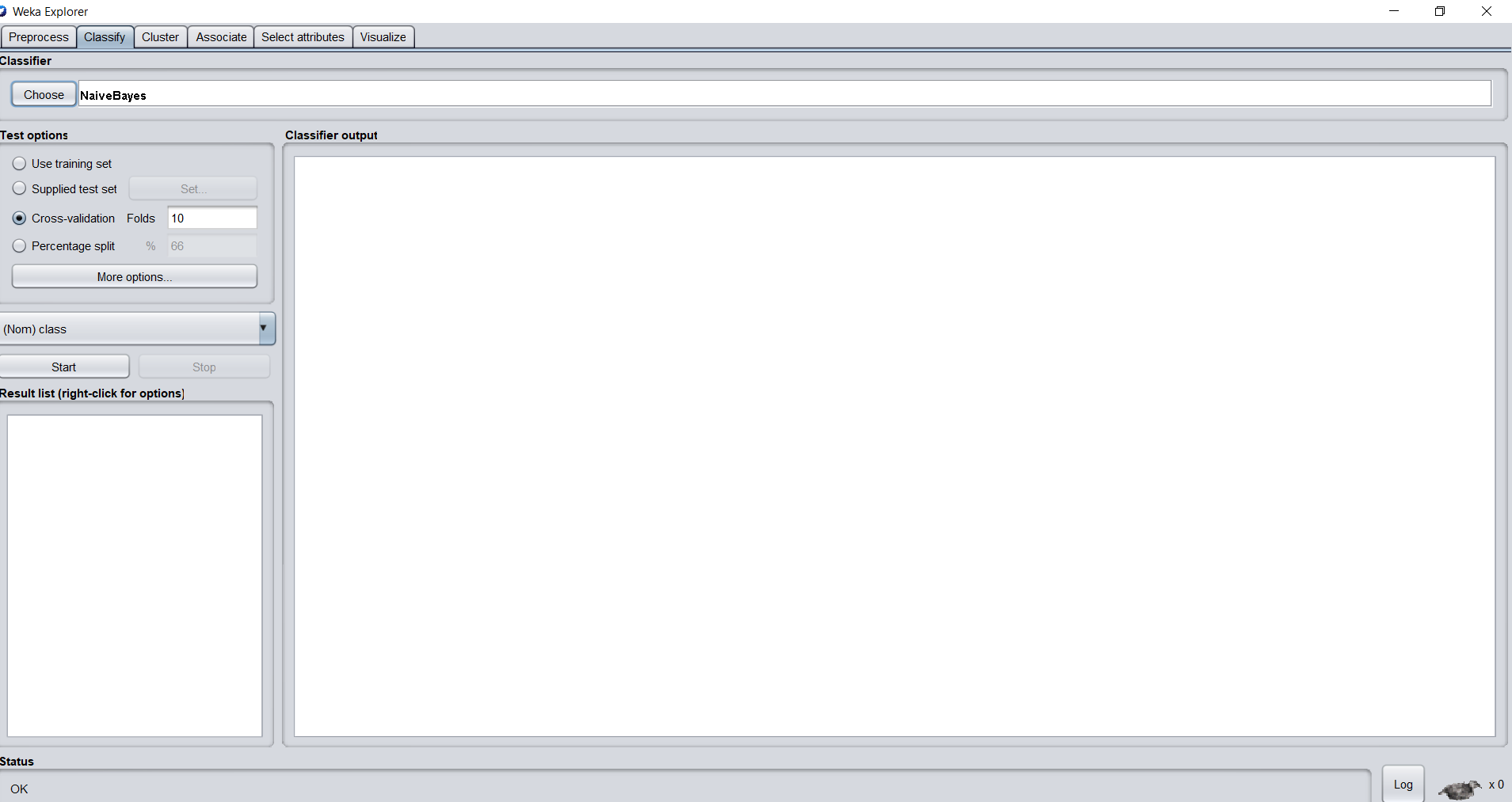
**SIT719 Security and Privacy Issues in Analytics**

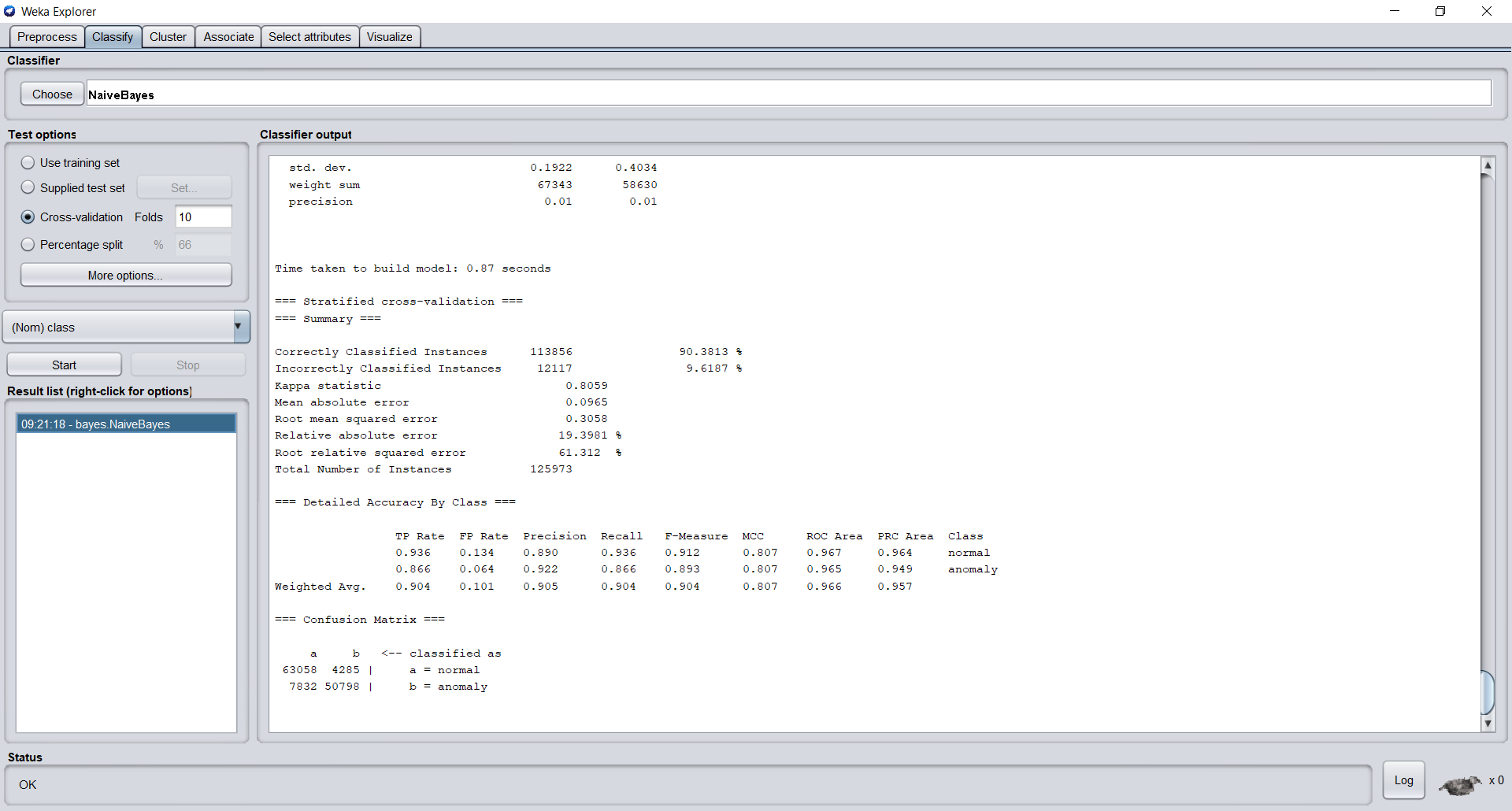
# Pass Task 4.1: Attack Classification using Naïve Bayes Algorithm

**Now apply “Naïve Bayes” classification algorithm from the “Classify” tab**

****

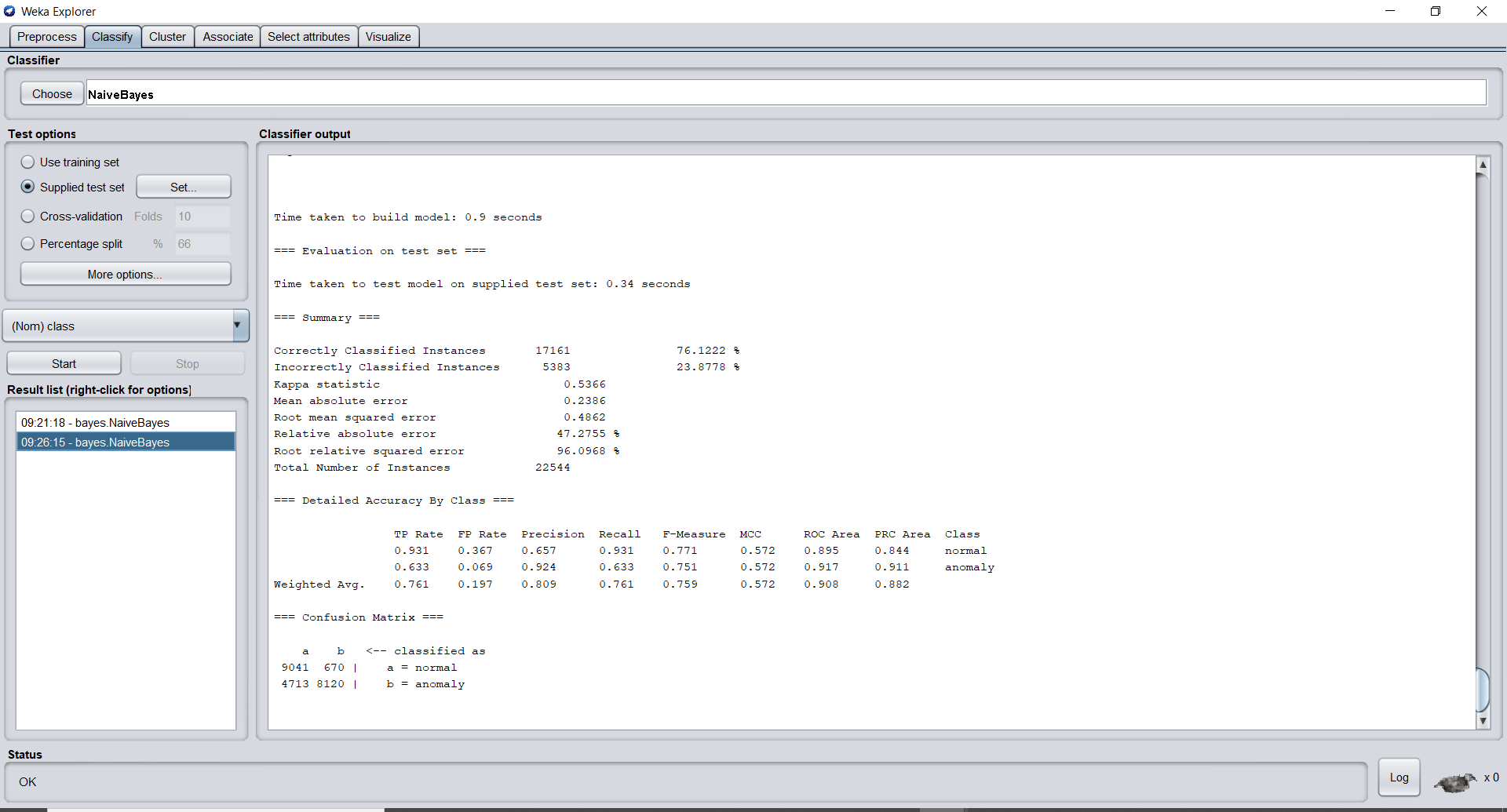
Successfully selected the“Naïve Bayes” classification algorithm from the “Classify” tab.

**Check the results with a 10-fold cross validation.**

****

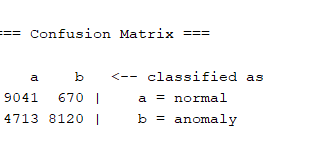
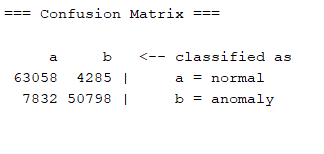
Successfully made the cross-validation with 10 folds and got the above results with correctly classified instances – 90.3813% and incorrectly classified instances – 9.6187%. Time taken to build the model is 0.87 seconds.

**Now, upload the test dataset and check the classification results**.



Successfully uploaded the test data and got the above results with correctly classified instances – 76.1222% and incorrectly classified instances – 23.8778%. Time taken to build the model is 0.9 seconds.

**Compare the results between 10-fold cross validation and the one obtained using the test dataset. Use confusion matrix to explain the results.**

****

**Confusion matrix for 10 folds Confusion matrix for test data**

On comparing both confusion matrix the matrix resulted using 10 folds has a high true positive value and which rated the correctly classified instances – 90.3813%. On comparing the false positive values, the matrix using 10 folds resulted the value of 4285 whereas the test data resulted 670. And the values of false negative are 7832 for 10 folds and 4713 for test data. The values of true negative are 50798 for 10 folds and 8120 for test data.

Finally, the output result obtained from the “Naïve Bayes” cross validation test option using 10 folds resulted in the high accuracy whereas the resulted output using test data is bit low accuracy on comparing to the cross validation.