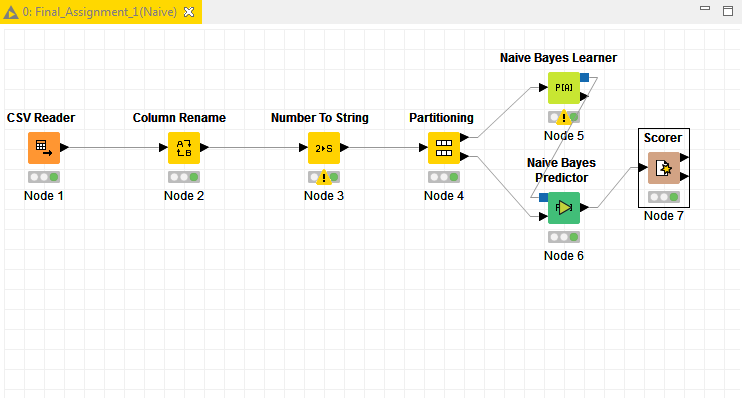
**Section 1(Supervised Learning)**

Here is the test result of 3 different classifiers:

1. **Naïve Bayes:**



**Summary:**

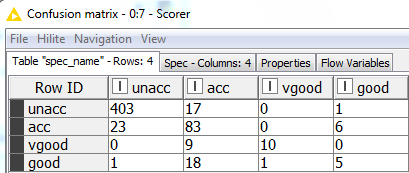
Accuracy = 86.828%

Error = 13.172%

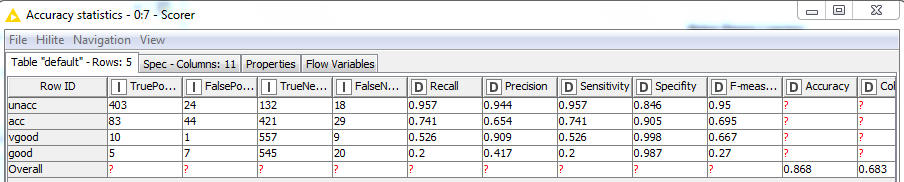
Correctly Classified = 501

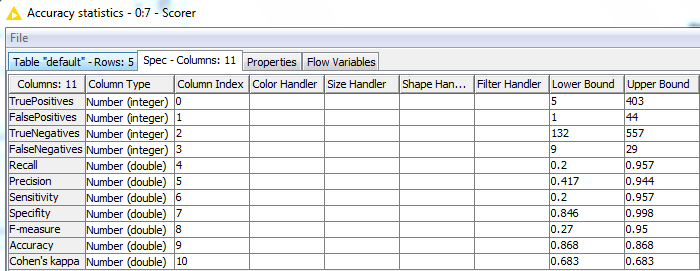
Incorrectly Classified = 76

Kappa statistic = 0.683

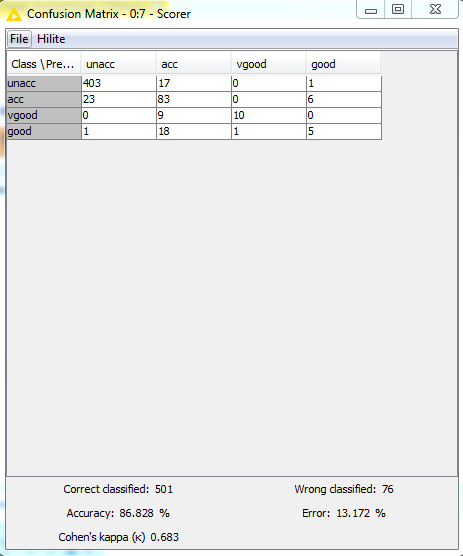
****

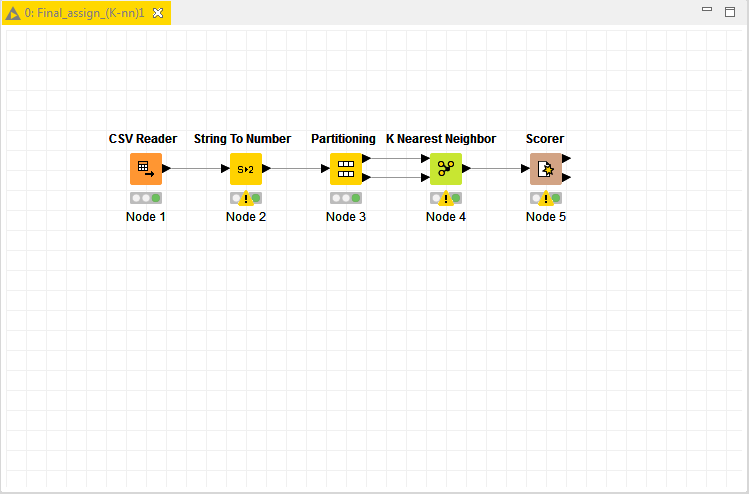
Confusion Matrix





Accuracy Statistics



1. **K Nearest Neighbor:**

**Summary:**

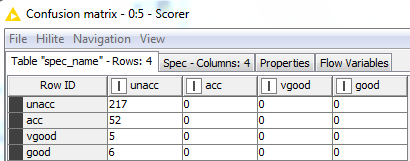
Accuracy = 77.5%

Error = 22.5%

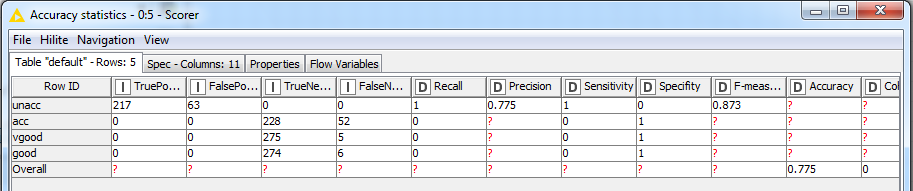
Correctly Classified = 217

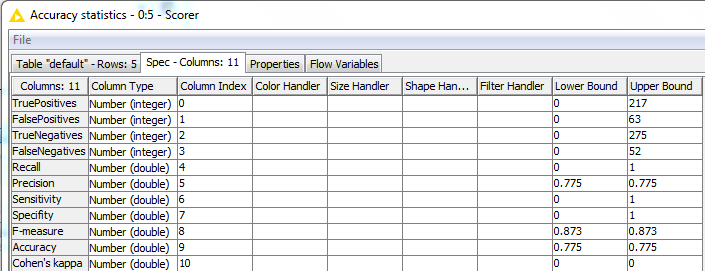
Incorrectly Classified = 63

Kappa statistic = 0.0

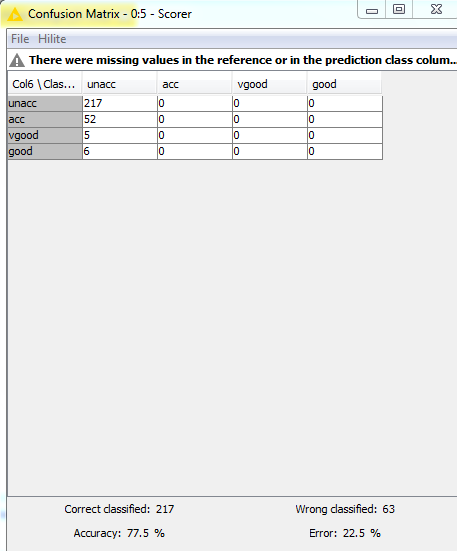


Confusion Matrix

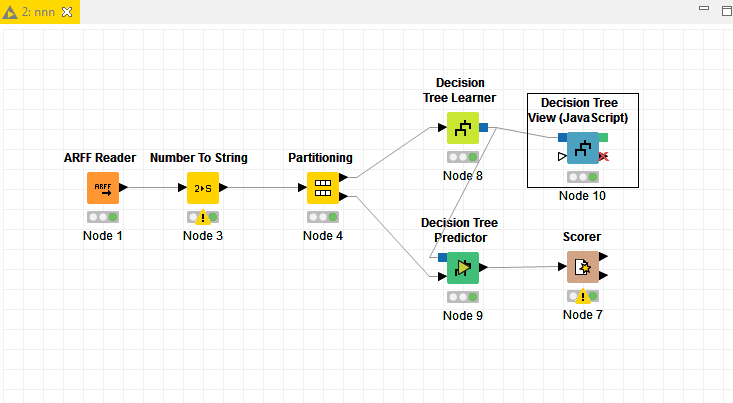




Accuracy Statistics



**Decision tree:**

****

**Summary:**

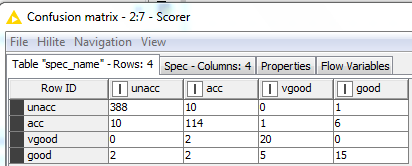
Accuracy = 93.229%

Error = 6.771%

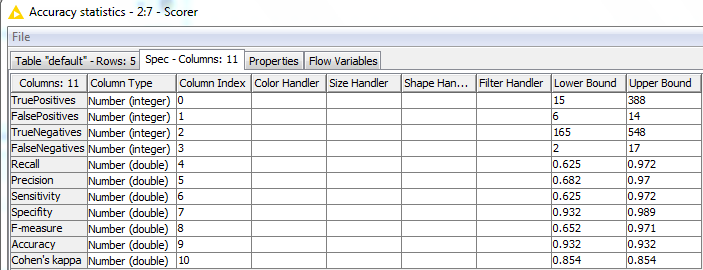
Correctly Classified = 537

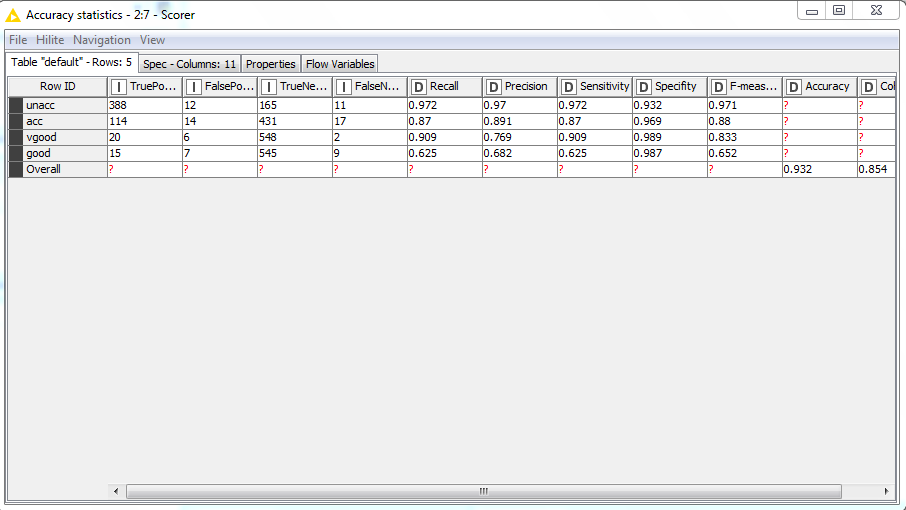
Incorrectly Classified = 39

Kappa statistic = 0.854

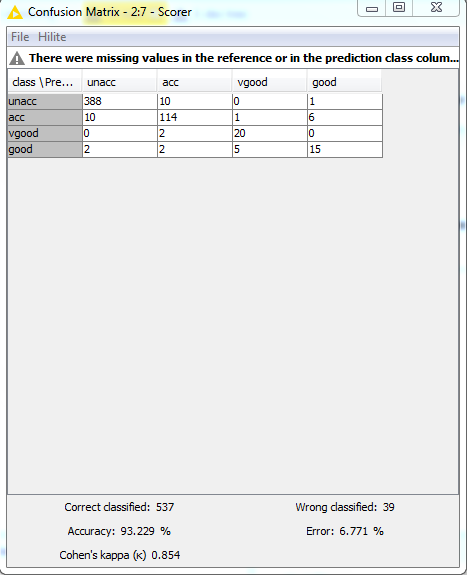


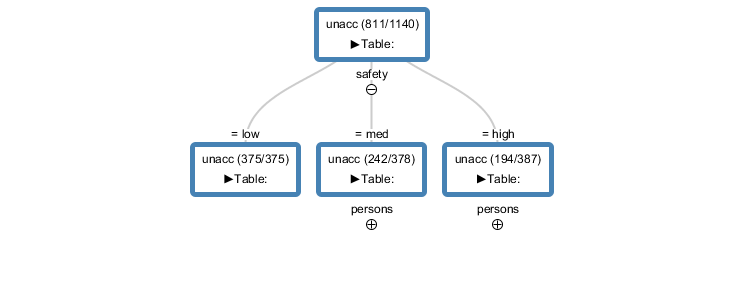
Confusion Matrix





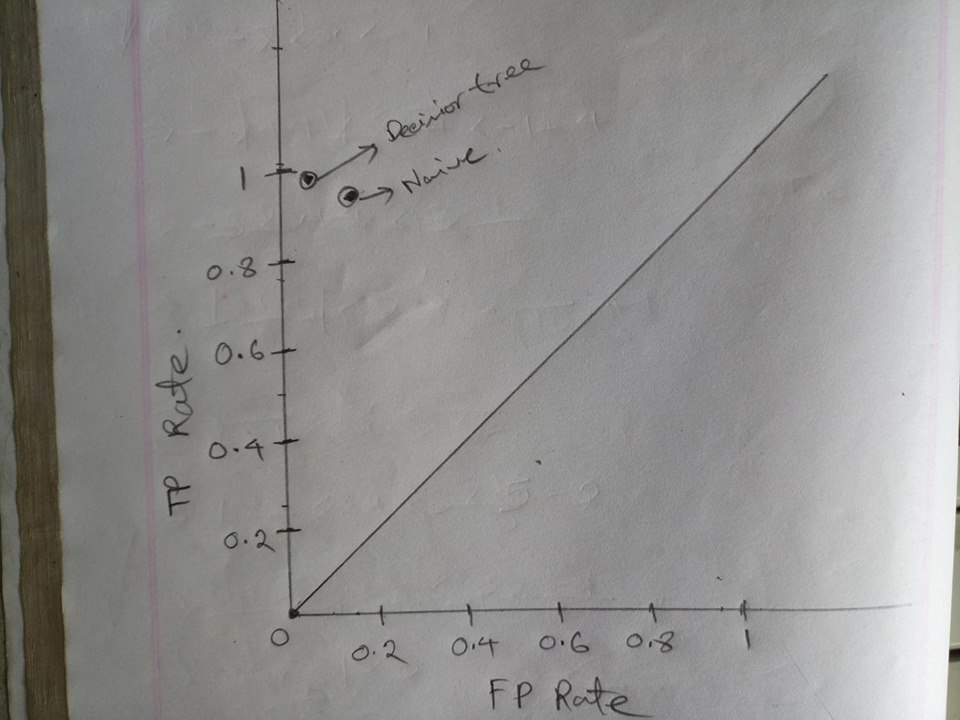
Accuracy Statistics





Tree View

**ROC graph:**



**Which classifier is suitable?**

For my point of view decision tree gives the best result among all correctly predicted classifiers and works extremely well on the training set. Its accuracy rate is more than 93%. Again by analyzing the ROC graph we can understand that decision tree is most optimum.

So I think Decision Tree is best suitable for this particular dataset.

|  |
| --- |
|  |