Part B

2. Compare the results of the three learning algorithms

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Accuracy | Precision | Recall | Time to construct model |
| Decision Tree | 90.65 % | 78.97 % | 69.34 % | 0.012177 Seconds |
| Gradient Boosting | 96.04 % | 78.37 % | 97.43 % | 0.0111412 Seconds |
| Random Forest | 89.35 % | 79.10 % | 64.89 % | 0.2964140 Seconds |

-200 words summary explaining how and why we would rank the quality of three models.

For how we rank the three models, we consider the time to construct models and the accuracy, precision and recall rate. As a result, the first is gradient boosting. And then the second is decision tree. The third is random forest.

We rank gradient boosting as first because it has the highest accuracy and recall rate, which are 96.04% and 97.43%. That are nearly 100%. The precision rate is 78.37%, which is almost the same as other two models. Also, it takes only around 0.01 second to construct gradient boosting model, which is the least among the three models. Therefore, we choose gradient boosting as the best quality model.

Then for the second good quality model, we compare the rest two models and choose decision tree as second. Although the accuracy, precision and recall rate of decision tree is overall same as the random forest, it takes only 0.012 second to construct model, which is less than the time to construct the random forest.

Finally we rank random forest as the third. The accuracy, precision and recall rate is a little bit lower than decision tree. And the time to construct model is a little bit more than decision tree.