

Fig – results of accuracies with k fold validation

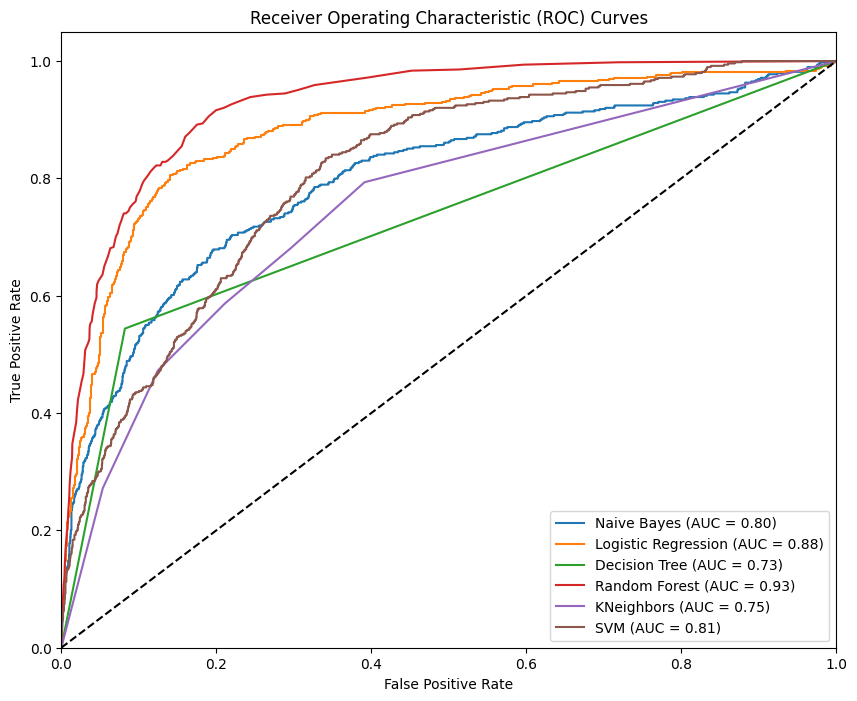


Fig – roc curves of all the models on the same plot

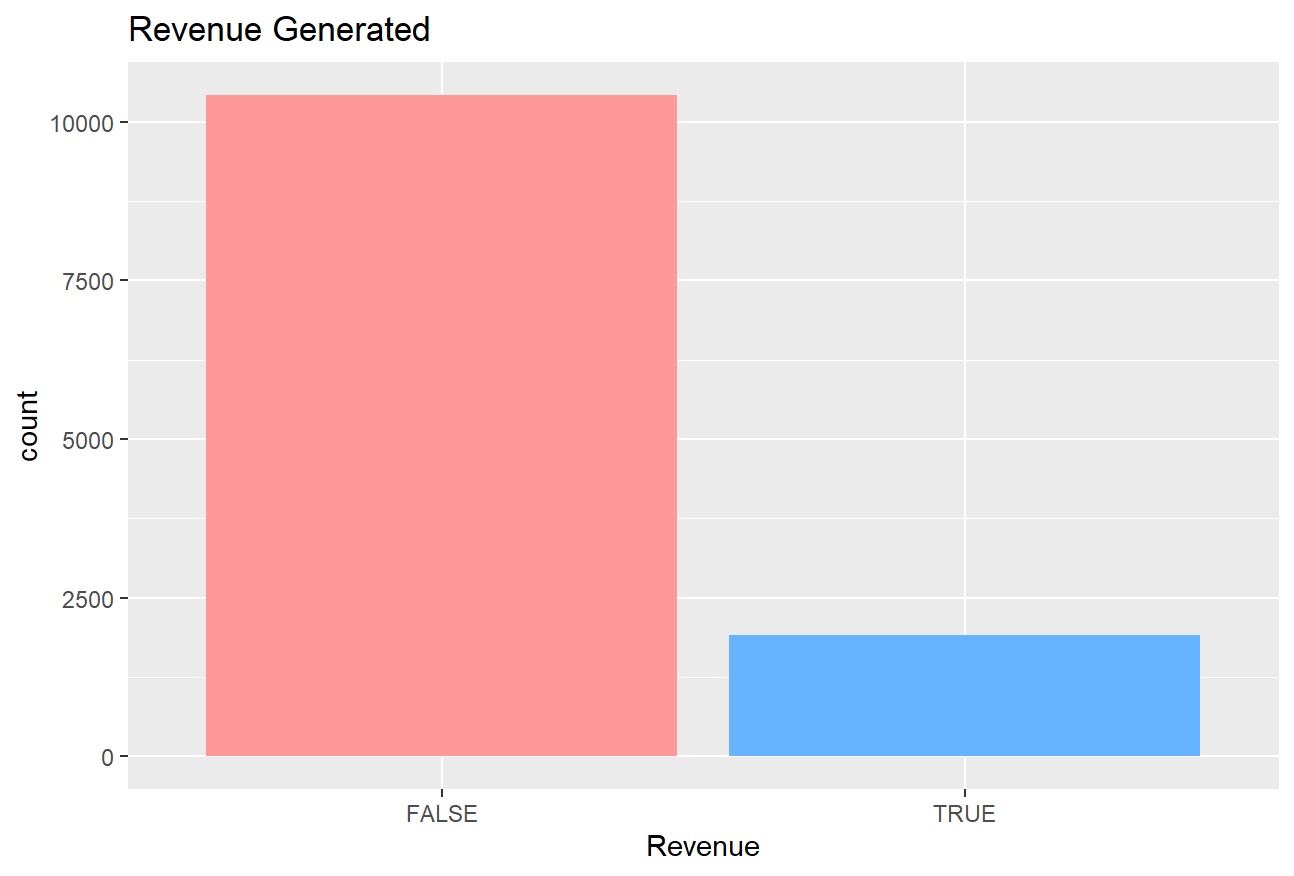


Fig – bar plot showing imbalance in dataset

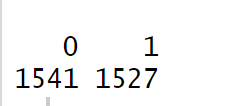
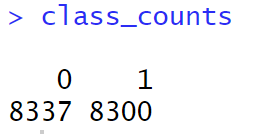
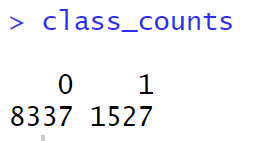


Fig – training datasets before and after smote and near miss algorithms

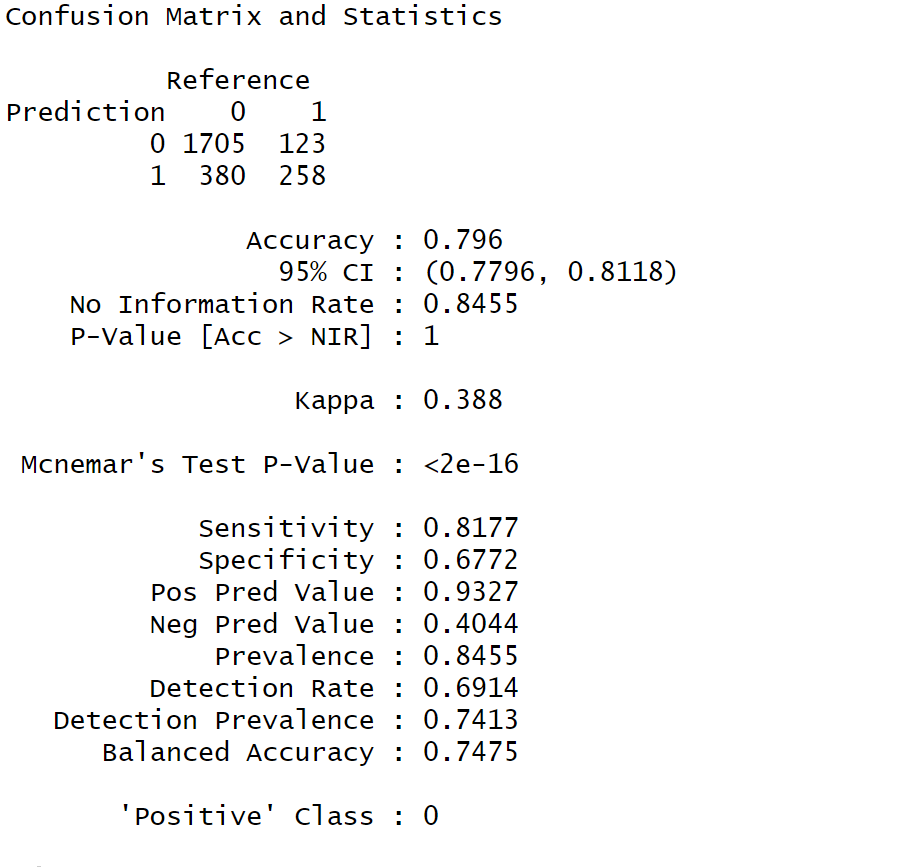


Fig – naïve bayes summary report

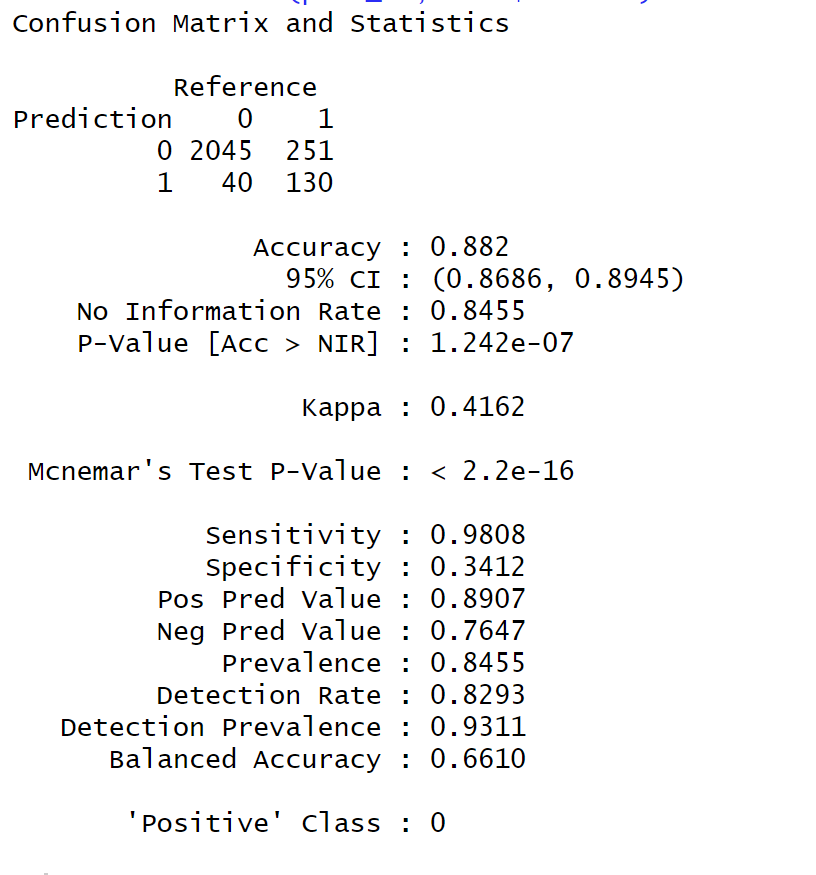


Fig – shows the summary report of the logistic model

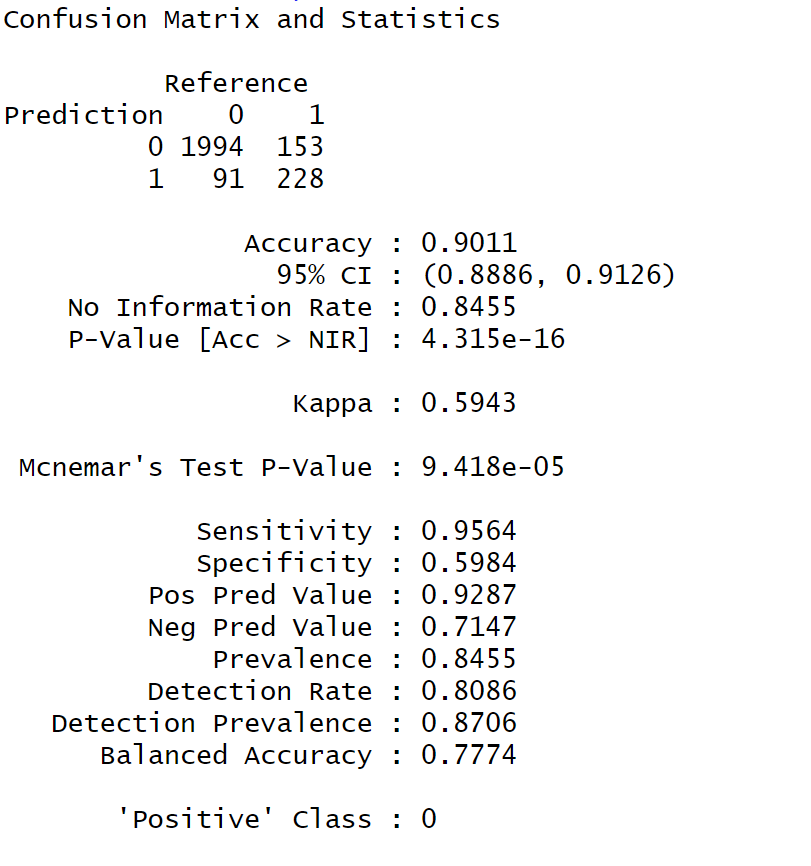


Fig – shows the summary report of decision tree model

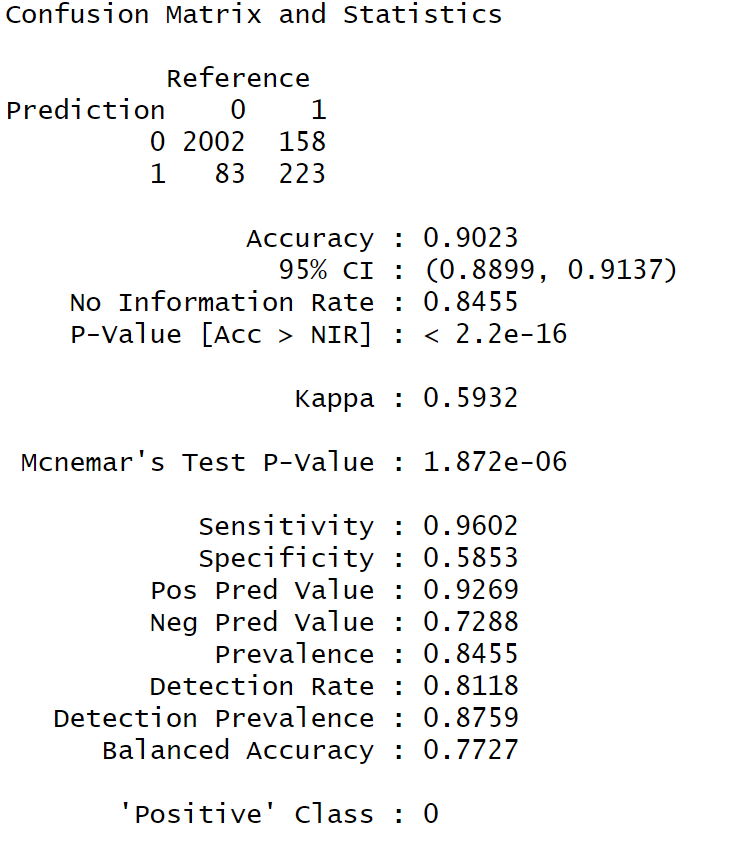


Fig – summary report of random forest

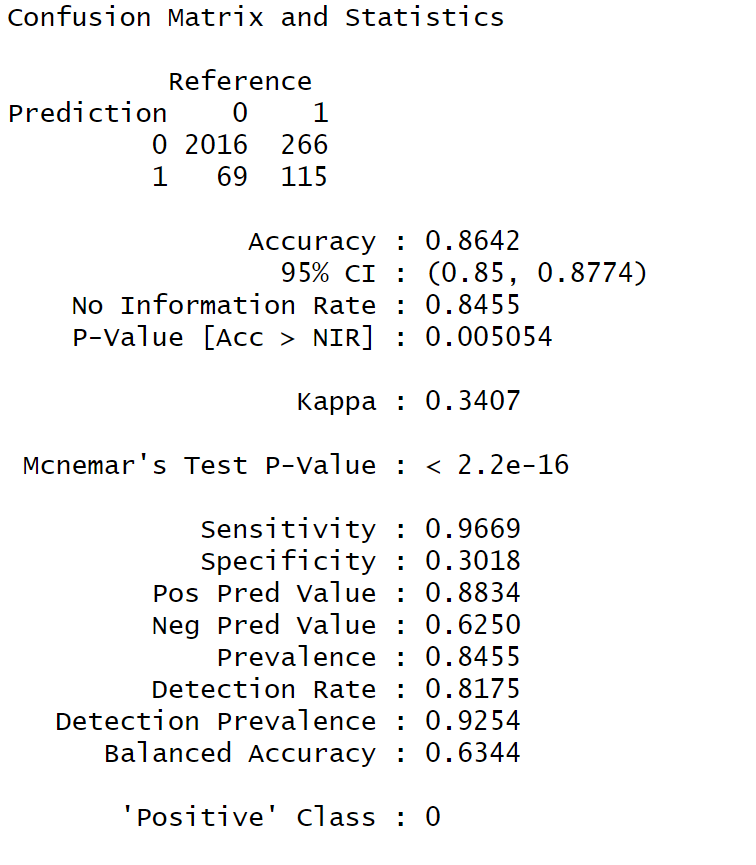


Fig – shows the summary report of KNN

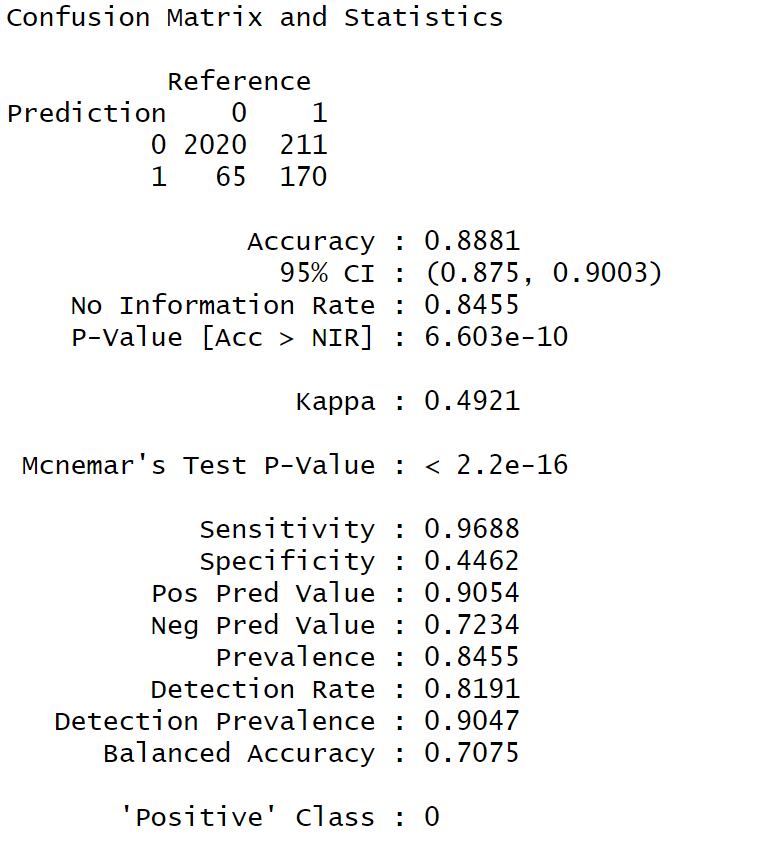


Fig – summary report of SVM

Similarly all the models shown above can be performed With an oversampled or an under sampled dataset.

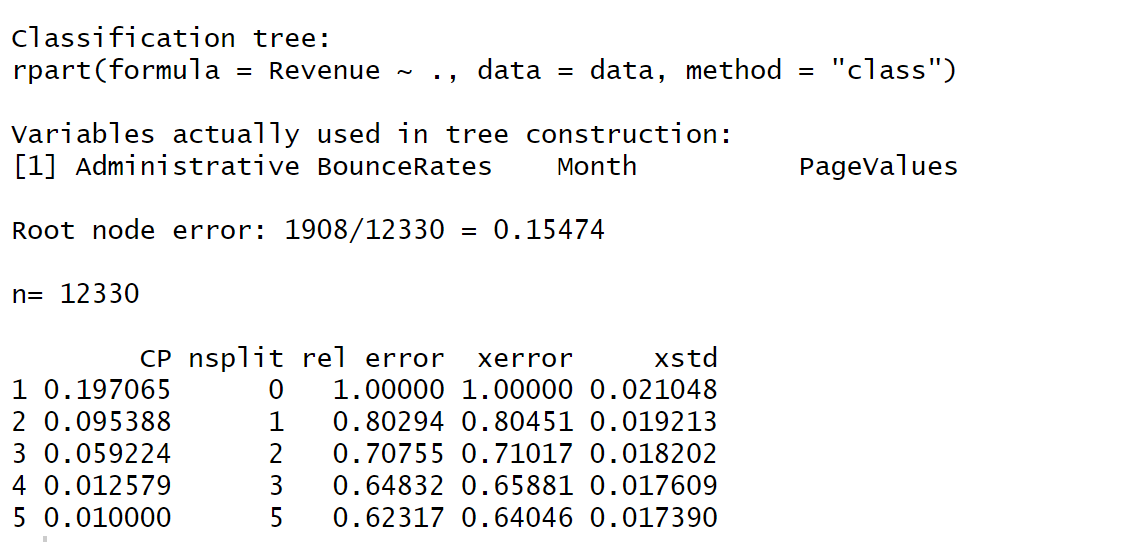


Fig – CP table for pruning in decision trees

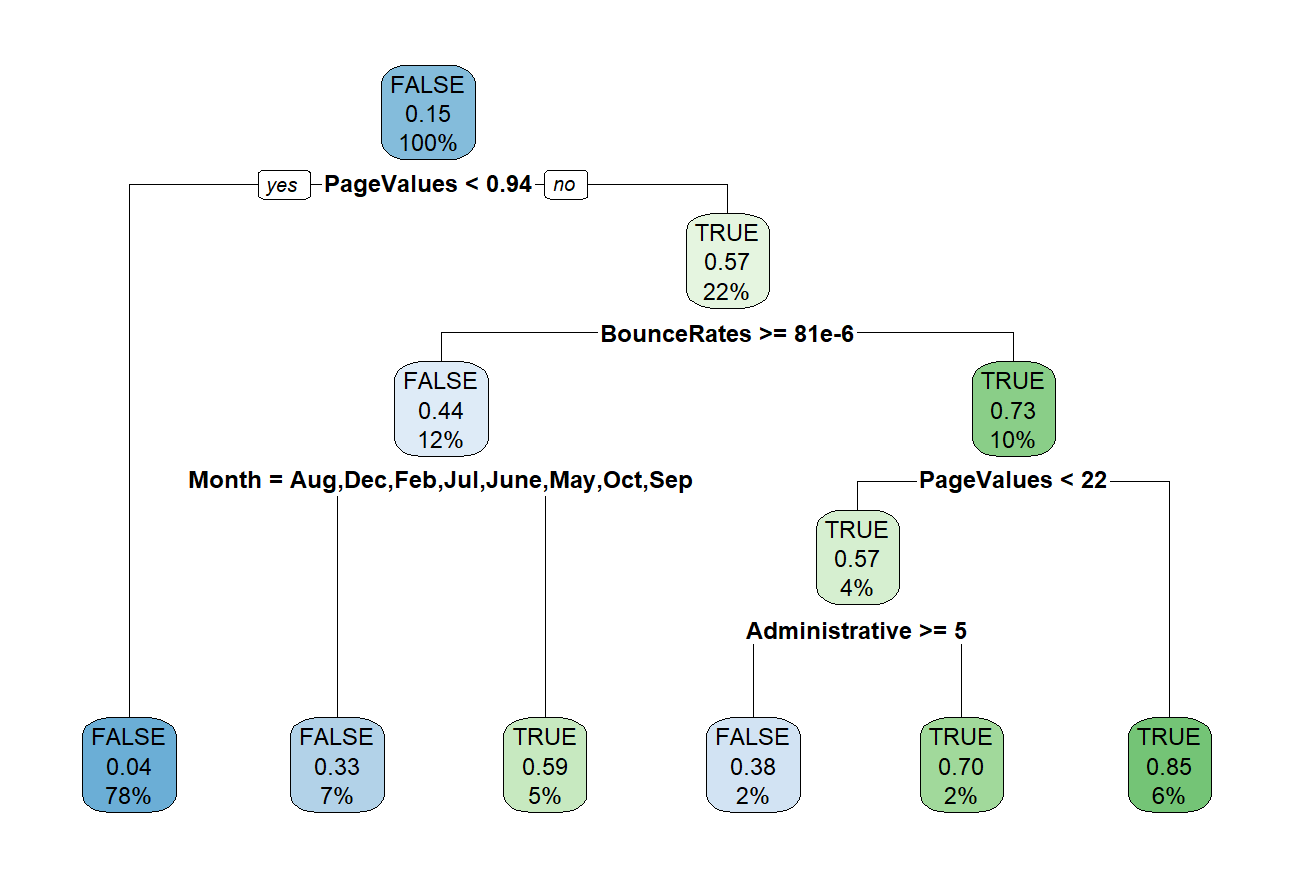


Fig – plotting what we have in had after the previous step

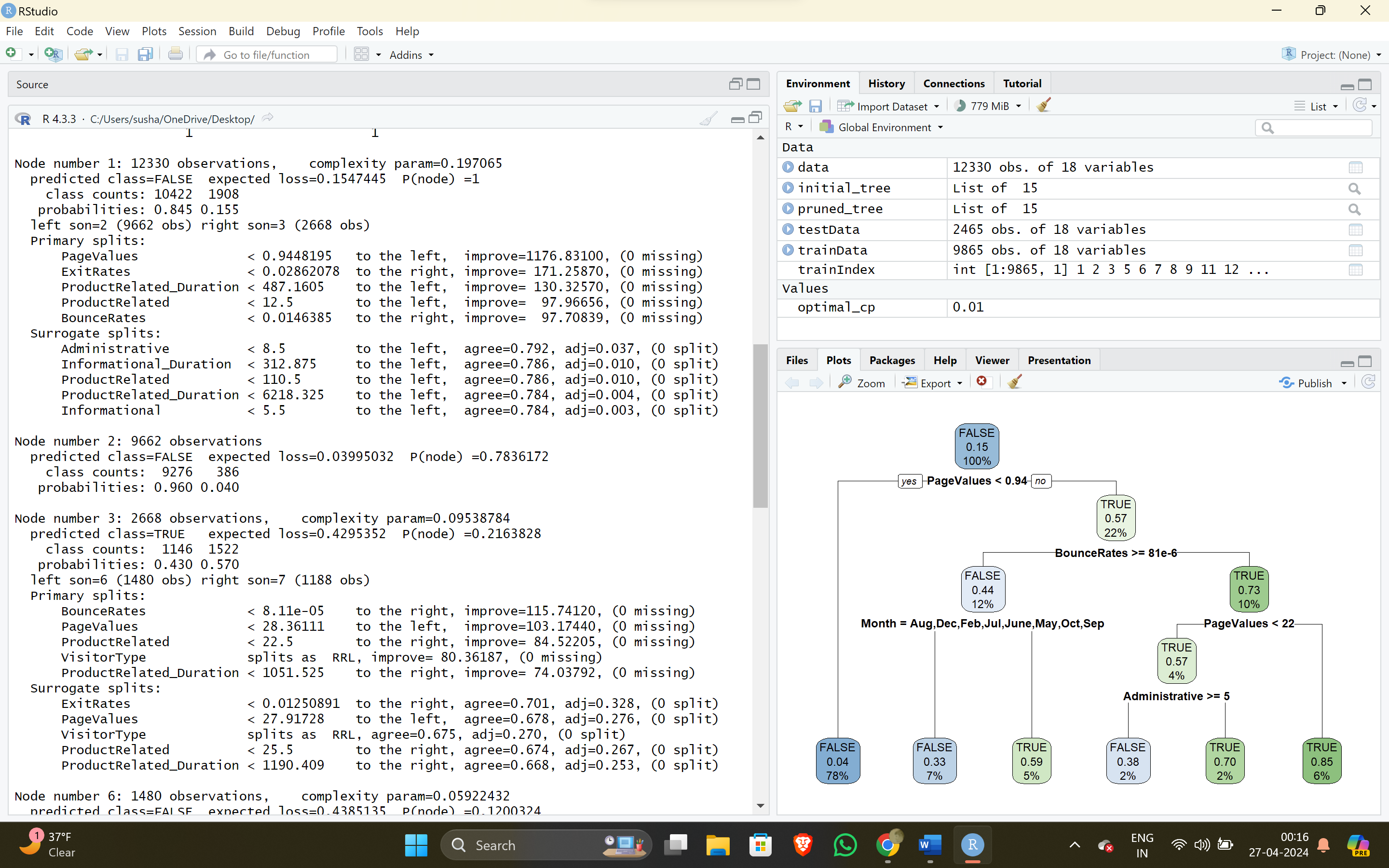


Fig – summary of pruned tree

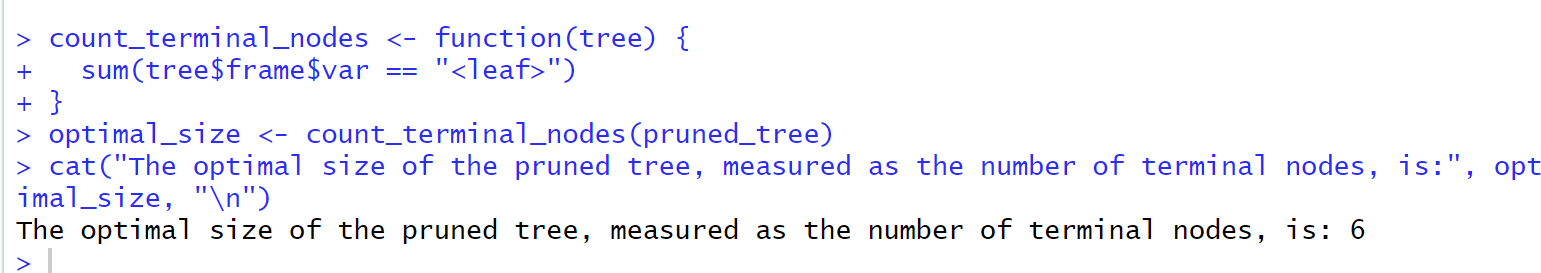


Fig – the size of the tree was found to be best at 6

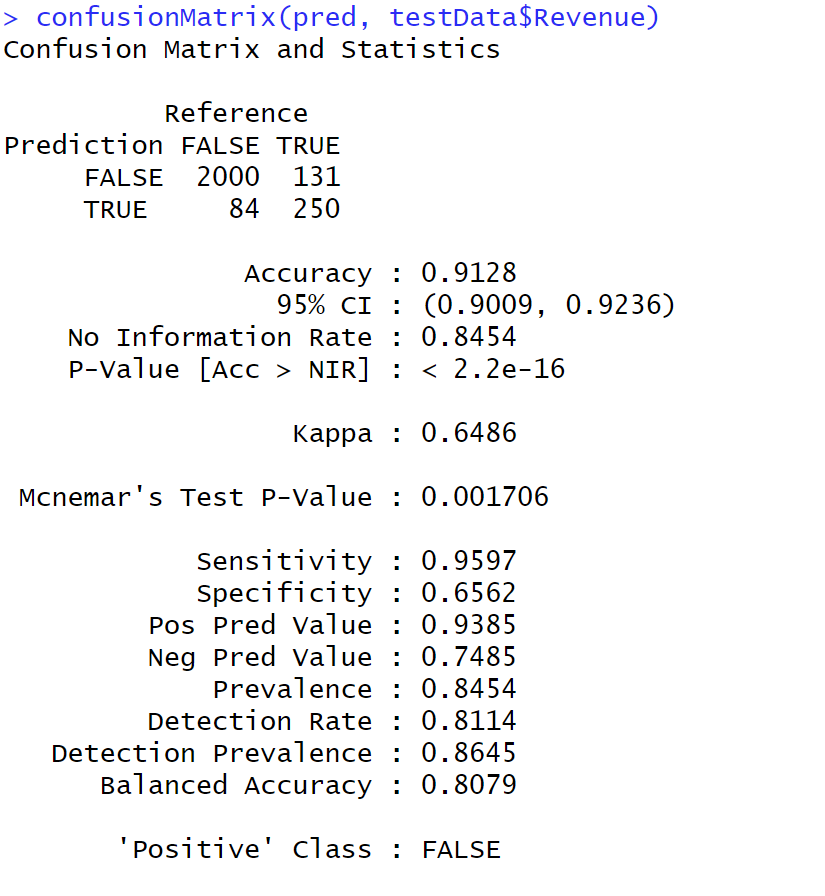


Fig - Decision tree prediction after pruning