LOGISTIC REGRESSION

CONFUSION MATRIX (AT = 0.5)

SENSITIVITY	SPECIFICITY	ACCURACY
0.7558	0.7553	0.7555

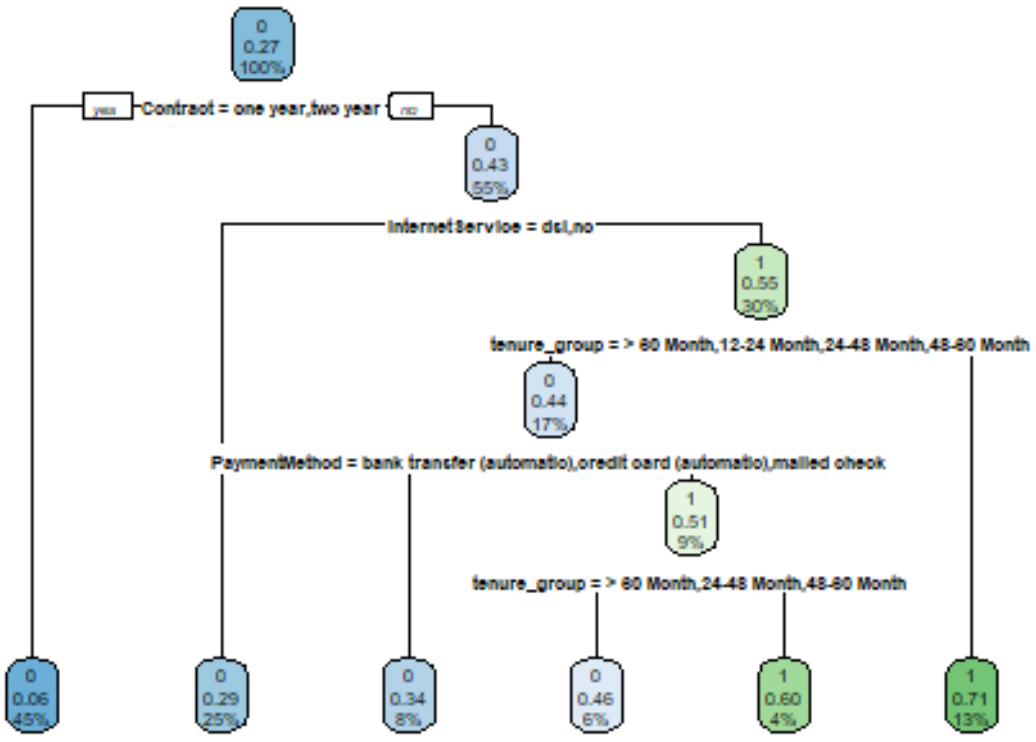
When the model is applied to the test data it yields accuracy of 79.57%, with 89% of customers that the model identified as leaving their service doing so and 51.87% of customers that the model identified as staying with the service staying with the company.

 # AT 0.3078247

SENSITIVITY	SPECIFICITY	ACCURACY
0.8128	0.7166	0.7422

Here sensitivity and specificity come closer to each other than before and accuracy remains same.

DECISION TREE



-the contract variable is the most important. Customers with month-to-month contracts are more likely to churn.

-Customers with DSL internet service are less likely to churn

- payment = bank-transfer are likely less to churn

- customer with greater than > 60 are less likely to churn

BEFORE TUNING ANY PARAMETER

SENSITIVITY	SPECIFICITY	ACCURACY
0.4260	0.9219	0.79

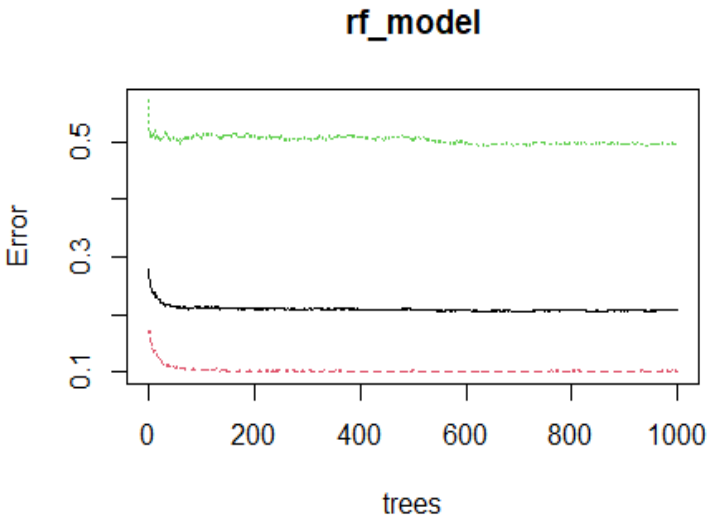
AFTER TUNING CP

SENSITIVITY	SPECIFICITY	ACCURACY
0.4528	0.8644	0.755

RANDOM FOREST

BEFORE TUNING ANY PARAMETER

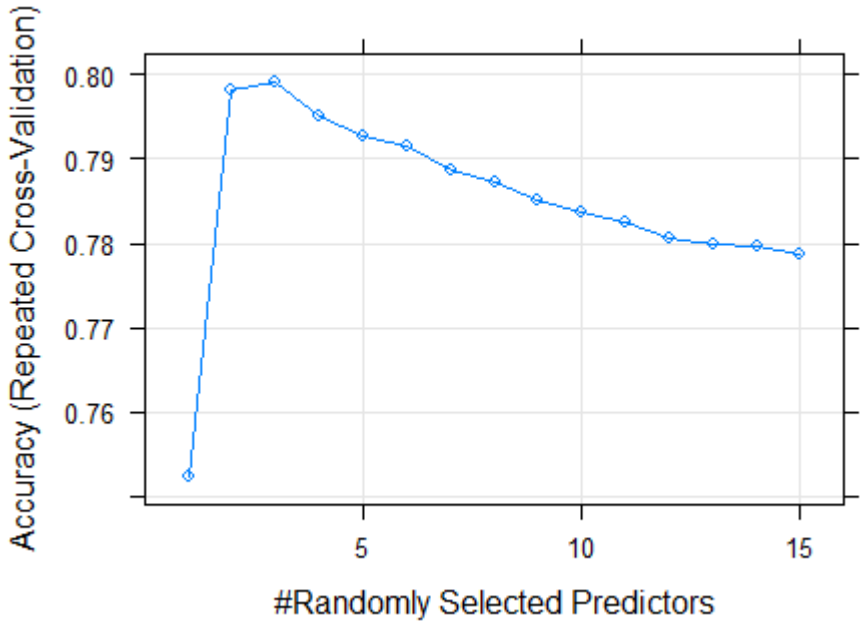
SENSITIVITY	SPECIFICITY	ACCURACY
0.8890	0.4902	0.7829



as the no. of trees increases, the OOB error rate decreases and then become constant.

AFTER TUNING PARAMETERS

SENSITIVITY	SPECIFICITY	ACCURACY
0.8922	0.5009	0.7882



at mtry = 3 we have maximum accuracy and tune the model accordingly , but get the almost same accuracy . BUT, OOB estimate of an error rate is decreased from 20.85% to 20.56%.

KNN

AT K = 1 (DEFAULT)

SENSITIVITY	SPECIFICITY	ACCURACY
0.6791	0.8928	0.836

AT K = 7

SENSITIVITY	SPECIFICITY	ACCURACY
0.7041	0.9367	0.8749

The KNN model has predictive accuracy of 87% and correctly classifies the 71 % are likely to churn.

SVM

SENSITIVITY	SPECIFICITY	ACCURACY
0.4902	0.9077	0.7967

The svm model gives accuracy of approximately 80% and 49% of customers are going to churn and correctly classified.

Naïve Bayes

SENSITIVITY	SPECIFICITY	ACCURACY
0.6865	0.8085	0.7761

The naïve bayes model gives accuracy of approximately 78% and 68% of customers are going to churn and correctly classified.

Thus from above prediction we observed that the Logistic Regression accuracy and positive results are higher as compared to other models.

Therefore it will be consider for further evaluation purpose.