

Axis Bank Machine Learning Intern Hiring Challenge - IIT Kharagpur

Saurav Kumar

15MT3FP11



AXIS BANK

Data-pre processing

1. **Replace** all the NA value with mode of that columns
2. convert the String columns into Integer using Label Encoder of sckit-learn package.

Model applied

1. Splitted the training data into train and test in 7:3 ratio.
2. Use Decision tree, gradient boosting with the default and modified hyperparameter(learning rate=0.01, no. Of tree=1000),Xgboost.

Final Accuracy(AUC)(on test data)

1. using decision tree = .52
2. Gradient boosting with the default hyperparameter= .75
3. Gradient boosting with modified hyperparameter = .7667
- 4.Xgboost = .74

Conclusion

Gradient boosting perform better the decision tree because it uses many decision trees and also boosting approach to train the model.

With modified hyper parameter , learning rate become small(large time to train) more appropriate prediction of weights. Also increase in no. Of decision trees from 100 to 1000 give better model training.