CS5590 – Foundation of Machine Learning

HACKATHON:

DRIVER FAULT CLASSIFICATION

Submitted By:

ANKITA JAIN (BM21MTECH14001)

TUSHAR K RAYSAD (BM21MTECH14004)

The models we tried for this Hackathon are:

1. ***Support Vector Machines*** : For this classifier we have used the hyper parameters C = 0.001, kernel = 'rbf', gamma = 0.01 and got the accuracy 0.55019.
2. ***Random Forest*** **:** For this classifier we have default values for hyper parameters and the accuracy is 0.84802.
3. ***Gradient Boosting*** **:** For this classifier we have grid searched for the hyper parameters n-estimators in the range(20,51,10), Learning rate in the range(0.01, 0.1 ,1 ,10), maximum depth in the range(5, 16, 5), minimum samples leaf in the range(30 ,71,10). And the accuracy is 0.85011 for the hyperparameters n_estimators=50, learning_rate=1, max_features=6, max_depth=15, min_samples_leaf=60, random_state=0.
4. ***XG Boost*** **:** For this classifier we have grid searched for the Hyperparameters min_child_weight in the range [1, 5, 10] , gamma in the range [0.5, 1, 1.5, 2, 5], subsample in the range [0.6, 0.8, 1.0], colsample_bytree in the range [0.6, 0.8, 1.0], max_depth in the range [3, 4, 5] and got the accuracy of 0.86309. This model gave the best accuracy.
5. ***Decision Tree*** **:** For this classifier, we have used hyperparameters such as max_depth = 20, min_samples_split = 500, max_features = 6, and we got accuracy 0.78879.

Conclusion :

As mentioned above, we got best accuracy in Gradient boosting and XGBoost classifier we have used those classifiers for final submission.