**Kaggle ID: 1438522**

I tried solving using various ensemble methods, these are the observations for different methods:

**Method 1: Gradient Boost Classifier:**

Using sklearn.GradientBoostingClassifier and with default hyper parameters I’ve got 0.9002 accuracy on the validation set (20% of training set)

On real test data, this model gave the accuracies Private Score: 0.90015 Public Score: 0.89983 on leaderboard.

**Method 2: Grid search for Gradient Boost Classifier to tune hyper-parameters:**

As grid search on whole training set will take time, I did grid search on 10% of training data set and finally optimized to

learning\_rate=0.1, n\_estimators=130, subsample=0.5, min\_samples\_leaf = 90, max\_depth=4, min\_samples\_split=2500

This model gave accuracy of 0.9019 on validation set and Private Score: 0.90198 Public Score: 0.90138 on leaderboard.

Method 3: Xtreme Gradient Boost Classifier:

With hyper parameters max\_depth=4, n\_estimators=140, learning\_rate=0.3, XGBoost classifier gave accuracy of 0.9032 on validation set and Private Score: 0.90309 Public Score: 0.90277 on leaderboard