

Prediction Report

Team: Habiba Shaik, Neelesh Nidadhavolu

File Structure:

- History
- Test
- Evaluation

History – Execution Time : 22 Minutes

- **Mapper** : Read all the a6history files in all directory and send only the required columns data to reducer with day as key.
- **Reducer** : Create Random Forest for each day and write ObjectOutputStream each Random Forest.

Test – Execution Time : 3 Minutes

- **Mapper** : Read the a6test file in all directory and send only the required columns data to reducer with day as key.
- **Reducer** : ObjectInputStream the required Random Forest and Random Forest classifies the output.

Evaluation – Execution Time : 30 Seconds

- Reads the a6validate file in all directory, compares it with Test output and returns accuracy.

Confusion Matrix

tp - 1319373.0
fp - 1164283.0
fn - 1265183.0
tn - 1635893.0
total - 5384732.0

Accuracy- 0.5488232283426547

Precision- 0.5312221177167853

Recall- 0.5104834253929882