Summer Internship Program – 2018

Problem Statement:

Given few attributes of voters, the objective is to come up with a model that predicts the "**religion**" of voter. Due to privacy, few of the attributes have been encoded.

To build this model, metric is 'accuracy'. You need to build a model where the system uses **train_dataset.csv** to train. After the training, make the predictions on the **test_dataset.csv** and share the CSV file along with predictions.

File descriptions:

- train_dataset.csv - Training set
- test_dataset.csv - Test set

Data fields:

Each row of the training data has the following features.

- Guardian_gender: gender of the voter's guardian.
- Guardian_name: name of the voter's guardian.
- Voter_age: age of voter
- Voter_gender: gender of voter
- Voter_name: name of voter
- age_as_on: date when the voter has cast the vote.
- Guardian_gender: gender of the voter's guardian.
- Guardian_name: name of the voter's guardian.
- Voter_age: age of voter
- Voter_gender: gender of voter
- Voter_name: name of voter
- age_as_on: date when the voter has cast the vote.
- House: address of voter(encoded)
- Year_of_Revision: revision year
- Assembly_Constituency_name: constituency name
- religion: religion of the voter
- index: miscellaneous

Attachments: Zip file with two data sets