

Readme file

Team Leader: Rahul Moorthy

Team Members:

1. Geet Shingi
2. Shahrukh Khan
3. Yakshit Jain

1. Preprocessing of dataset :

- a. Splitting date time feature into different components like **day,month,year,week of year**,etc
- b. Writing to **is_holiday** column where value is not equal to None as the Holiday value is defined only for one particular time.
- c. Dropping the **outliers** row by analyzing the rain,snow and temperature values.
- d. Dropping columns like
 - i. Visibility
 - ii. Air pollution index
 - iii. Dew point
 - iv. Humidity
 - v. wind_speed
 - vi. Wind_direction
 - vii. Rain_p_h
 - viii. Snow_p_h
 - ix. Clouds_all
 - x. year
- e. Introducing **Peak time** as a feature by analyzing the hour and traffic volume features.
- f. One Hot Encoding Weather type and ranking it according to average traffic
- g. Label encoding the **holiday** feature by analyzing the traffic volume with respect to holiday feature.

2. Model trained :

`RandomForestRegressor(bootstrap=True,n_estimators=100)`

Files in the Folder

1. Data Exploration
2. Preprocessing and Final Model
3. Submission File