**An Intelligent Traffic Congestion Prediction using Machine Learning**

**Data Download instructions:**

The data is originally sourced from BigQuery-Geotab Intersection Congestion Data: <https://www.kaggle.com/competitions/bigquery-geotab-intersection-congestion/data>

The data is present in google drive with the name ‘train.csv’. The link is <https://drive.google.com/file/d/1IB1lMoRYQ3g--cFmGx6UEvmLN3YSyWUP/view?usp=share_link>. However, the code will automatically download the data from gdrive and perform all the operations. No action needed.

**The folder should have the below files**

* Data\_Preprocess.py
* gdrive\_downloader.py
* boost\_xg.py
* CatBoost.py
* LightGBM.py
* randomforest.py
* kfold\_CB\_model.pkl
* kfold\_LightGBM\_model.pkl
* kfold\_RF\_model.pkl
* kfold\_XG\_model.pkl
* main.py

**Install below libraries in the pycharm terminal/ before running the code**

* pip install pandas
* pip install sns
* pip install matplotlib
* pip install scikit-learn
* pip install gdown
* pip install lightgbm
* pip install xgboost
* pip install catboost

**Run the main.py file in the terminal**

mode : choices=['train', 'test']

model : choices=['LightGBM', 'CatBoost', 'XGBoost', 'RandomForest']

**Run respective commands to see the results:**

python main.py --mode train --model LightGBM

python main.py --mode test --model LightGBM

python main.py --mode train --model XGBoost

python main.py --mode test --model XGBoost

python main.py --mode train --model CatBoost

python main.py --mode test --model CatBoost

python main.py --mode train --model RandomForest

python main.py --mode test --model RandomForest

**Team Contribution:**

| **Tasks** | **Team Member** |
| --- | --- |
| Brainstorming and Topic selection | All |
| EDA | All |
| Feature Engineering | All |
| Random Forests | Divya |
| XGBoost | Divya |
| CatBoost | Nupur |
| LightGBM | Vani |
| Report | All |
| Presentation | All |
| Coding Standards | All |