**Project 1015 – Taxi Fare Prediction system.**

**The project is planned to kick-off on 7th September 2021.**

Planned End Date: 15th Oct 2021 (Final review on this date).

Bharat Taxi is one of the cab services in India which is growing rapidly online car rental and taxi hire companies in India

**Summary-**Predict the optimized fare price: The system should be able to predict the competitive fare price depending on several factors. For example, during the peak hours, the prices of the fare should be comparatively high or if the travel is at night, then the night fares might apply. Other factors to consider are distance, weather, availability of cabs, etc. If the car is rented for >1 day, then the fare should be calculated accordingly. .

**Technology Stack:**

● Python 3x and Anaconda distribution (https://www.anaconda.com/products/individual)

● Jupyter IDE notebook /Jupyter Lab (comes with Anaconda distribution), access to Google Colab

● Python IDE (Atom, Eclipse etc)

● Plotting/visualization libraries: Matplotlib and Seaborn (again comes with Anacondadistribution)

● Machine learning libraries: Sklearn (again comes with Anaconda distribution), scipy etc.

**Datasets:**

Yellow/Green/For hire taxi-2017,2018,2019

<https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page>

A NOTE ON SHARED TRIPS In the 2018 FHV records, there is a field called SR\_Flag, which is supposed to indicate if the trip was a part of a shared ride chain offered by a High Volume FHS company (e.g. Uber Pool, Lyft Line). For shared trips, the value is 1. For non-shared rides, this field is null.

NOTE: For most High Volume FHS companies, only shared rides that were requested AND matched to another shared-ride request over the course of the journey are flagged. However, Lyft (hvfhs\_license\_num=’HV0005’) also flags rides for which a shared ride was requested but another passenger was not successfully matched to share the trip—therefore, trips records with SR\_Flag=1 from those two bases could indicate EITHER a trip in a shared trip chain OR a trip for which a shared ride was requested but never matched. Users should anticipate an overcount of successfully shared trips completed by Lift.