Project Proposal

Uber and lyft Price Prediction

Introduction

Sometimes it is just convenient to leave the driving to someone else. This is just less stress, more mind space and one gets to use that time to do other things. Well, that is one of the ideas that grew and later became the concept behind Uber and Lyft.

Both these companies offer passenger boarding services that allow users to rent cars with drivers through websites or mobile apps. Whether it is a short distance ride or you are going from one city to the other, these services have helped people in many ways and have actually made their lives lesser complicated.

Problem:

I will analyze to understand the factors affecting the dynamic pricing and the difference between Uber and Lyft's special prices.

Dataset:

- The dataset contains Uber & Lyft trip from 11-26-2018 to 12-18-2018 , which I will analyze to understand the factors affecting the dynamic pricing and the difference between Uber and Lyft's special prices.
- The datasets used in this project have been imported from Kaggle with two files. The First is cab_rides which contains 693071 rows× 8 columns, and the Second is weather contains 6276 rows × 8 columns
- The data include the following Features:

Cab_rides.csv
- **Distance**: The Distance Between Source and Destination

- **Cab_type**: Uber or Lyft
- **Time_stamp**: Epoch Time When Data was Queried
- **Destination**: Destination of the Ride
- **Source**: The Starting Point of the Ride
- **Price**: Price Estimate for the Ride in USD
- **Surge_multiplier**: The Multiplier by Which Price was Increased, Default 1
- **Id**: Unique Identifier
- **Product_id**: Uber/Lyft Identifier for Cab-type
- **Name**: Visible Type of the Cab Eg: Uber Pool, Uberxl

Weathr.csv

- **Temp:** Temperature in F
- **Location:** Location Name
- **Clouds:** Clouds
- **Pressure**: Pressure in MB
- **Rain**: Rain in Inches for the Last Hr.
- **Time_stamp**: Epoch Time When Row Data was Collected
- **Humidity**: Humidity in %
- Wind: Wind Speed in MPH

Tools

I will use :
- Environment: Jupyter Notebook.
- Programing Language: Python.
- Libraries: Pandas,NumPy,Seaborn ,Matplotlib ,seklearn
- Other: Machine Learning Algorithms (Linear Regression for modeling)

Solution

In this project, I will analyze to understand the factors affecting the dynamic pricing and the difference between Uber and Lyft's special prices. Using Machine Learning Algorithms to predict the determining factors that lead to surge pricing .