Project Proposal

Oct 4th 2020

Problem

Ride sharing has been one of the major commute options. For price sensitive users, the predicted price of each ride might influence their choice of ride sharing providers. However, the algorithm for pricing calculation is a blackbox for consumers. We then formulate our question: What are potential factors that influence the price? What company offers cheaper rides?

Dataset

We found a dataset consisting 693,071 rides, consisting 56% Uber rides and 44% Lyft rides, during November and December in Boston, MA. Each ride comes with information such as the day and time when it was scheduled, the distance of the ride, the kind of cab (normal, XL, Lux, etc.), the temperature and precipitation probability during that time, and the price of this ride. We plan on using all these variables to construct a model to predict the price of a ride for both Uber and Lyft.

Data from Kaggle.com:
`https://www.kaggle.com/brllrb/uber-and-lyft-dataset-boston-ma`

Value

The proposed project will allow us to predict the price of both Uber rides and Lyft rides with given values on parameters such as time, distance, and weather. Individuals can use the result of the project to better plan their schedules, choosing the time and the product that would save the most cost for example. It would also be interesting to see the price comparison between Uber and Lyft with the same factors, giving a reliable first choice for individuals based on price. So instead of opening both applications, typing in the destination twice, and comparing the price between the two, the individual can have a price expectation in mind and use the one that generally gives a lower price.