Analysis on supply and demand gap in transport to and from airport using uber

Abstract— To analyze supply and demand gap in transport to and from airport using Uber sample data, as this data is being generated in real time in huge chunks I am using Hadoop for analyzing this data.

INTRODUCTION

You may have some experience of travelling to and from the airport. Have you ever used Uber or any other cab service for this travel? Did you at any time face the problem of cancellation by the driver or non-availability of cars?

Well, if these are the problems faced by customers, these very issues also impact the business of Uber. If drivers cancel the request of riders or if cars are unavailable, Uber loses out on its revenue. Uber is facing – driver cancellation and non-availability of cars leading to loss of potential revenue.

The aim of analysis is to identify the root cause of the problem (i. e. cancellation and non-availability of cars) and recommend ways to improve the situation. As a result of your analysis, you should be able to present to the client the root causes and possible hypotheses of the problems and recommend ways to improve them.

UNSUPERVISED METHOD

Uber or any cab service is facing a low revenue cost due to non availability of cabs in morning hours from city to airport and from airport to city in evening hours. In the first scenario that is from city to airport in morning hours, the cab drivers are mostly in the city due to more number of customers in the city. These customers are mainly employees using cab service to reach their

workplace, students trying to reach their colleges etc. So the cab driver would not like to waste the fuel by coming alone from airport. Whereas in the second scenario that is from airport to city in the evening hours, the cab drivers would be waiting for the arrival of flights and would not like to waste the fuel by travelling alone at the time of return from the city.

DISADVANTAGES:

LOSS OF REVENUES TO UBER: Due to cancellation of cabs, the revenue to uber gets low.

SURGE PRICING: Surge pricing is to adjust prices of rides to match driver supply to rider demand at any given time. During periods of excessive demand when there are many more riders than drivers, or when there aren't enough drivers on the road and customer wait times are long, Uber increases its normal fares. They do this with a " multiplier" whose value depends on scarcity of available drivers. This affect the customers.

SUPERVISED METHOD

To overcome this issue we need to collect all bookings data of uber. Our normal systems can allow gigabyte of data. But the data at uber is in terabyte. As the data is in terabyte we use Hadoop and Spark frameworks to allow distributed processing. We also use data cleaning and data manipulation techniques to avoid null values and also to manipulate some data to analyse the exact issue of the problem faced by customers and the company.

Data Manipulation is done by adding extra columns like date, hours, and days to the existing data. Date column is added to change the format of date type. The hours column in the data is of exact timestamp that includes hours, minutes, seconds, milliseconds. This timestamp is of no use to us we just require hours and minute so we change this time timestamp to just hours and minutes by adding an extra column. Days column is created to segregate the given data in days of the week. All these manipulations are done to clearly analyse the data and map it using ggplot.

Data Cleaning is done to avoid null values that we get from cancellation of bookings from the customers.

After data cleaning and data manipulation we map the data altogether using ggplot function from the library of RStudio. We get the graphs of the bookings based on hours, dates and days.

RESULT:

Problem due to unavailability and cancellation has been solved.

Revenue can be increased.

Number of trips in morning are high from the city.

Number of trips from the evening are high from the airport.

Major problems are:

Cancelled trips during the morning rush 2. Unavailability of cars during evening rush

CONCLUSION

For the trips in the morning, drivers can be incentivized to make those trips.

They could be given a bonus for each trip they complete from the city to the airport in the morning rush. This will ensure that less number of trips are cancelled.

Uber can pay for the gas mileage of drivers to come back to the city without a ride.

Uber can increase the demand at the airport to reduce idle time by increased marketing and price cuts for the passengers.

For the evening, since the number of drivers is less, some of the ways are:

Drivers can again be given a bonus to complete a trip from the airport in the evening. This will ensure that the supply increases at the airport.

Uber can also pay drivers to come without a passenger to the airport.

Another innovative way can be to pool the rides of passengers so that lesser number of cars can serve more passengers