Project Proposal

Student Name: - Jitendra Agarwal

Course: -   Springboard cohort Jan2

Project: -   Capstone2

Proposal: -   Citi Bike, NYC – Repricing case study and Twitter/Yelp Sentiment analysis for real time reputation management

Client: -   Citi bike NYC- Leading Bike sharing program in US

Summary: -   Citi Bike is the largest bike share program in us, with 10,000 bikes and 600 stations across Manhattan, Brooklyn, Queens and Jersey City. It was designed for quick trips with convenience in mind, and it's a fun and affordable way to get around town. Everyone knows that bike sharing is the answer to many environmental and urban transportation issues, yet it's not mainstream in US.

I am being asked by the senior executive team at Citi bike to use data science techniques to recommend 3 key action item to increase the company's' business.

Problem Statement: -

- Citi bike management is curious to know if there is any statistical way to find the reason in decline in trips observed few times last quarter.
- How are user sentiments about a recent change in pricing and if Citi bike is really useful for users from time, cost and efficiency point of view compared to other transportation options?
- How is the performance of unit economics per trip or per bike or per station?
- What is the most common use of Citi bike?
- What measure can be taken to increase user trips by 5% with existing users.

Project Goals: -

- User Sentiment analysis from twitter and identify most common customer issues and how to address them.
- Analyze if any of the current plan can be repriced to get 5% increase on revenue with no customer impact.
- Can we predict a right number of bike to stationed as a station?
- Propose a new monthly pass pricing for office commuters.

Approach: -

- Business use case analysis and EDA for unit economics performance
- Clustering and Geospatial analysis
- Time series and Network analysis
- Regression and Machine learning models for predictions
- NLP -User sentiment analysis based on twitter hashtags and yelp reviews

Data: -

- Bike trip data provided by Citi bike: https://www.citibikenyc.com/system-data
  We will use Q4 2017 data for this analysis
- Daily weather data from open sources

Deliverables: -

Source code in python notebook on github repository
Project Summary report for client with recommendation
Technical design document
Blog post