BELLEVUE UNIVERSITY

Final Project Proposal

DSC630 - Prodictive Analytics

Gangadhar Dhulipala
Anirban Pal
Raghu Raman Nanduri
June 28, 2019

Contents

INTRODUCTION

1 PROJECT BACKGROUND

The project is a part of the final submission for the Summer 2019 Predictive Analytics class (DSC630) as part of the MS Data Science program in Bellevue University, Nebraska. We would be predicting the outcome of a specific scenario, based on open data available in the internet, by applying predictive modelling and other machine learning concepts.

Dataset Name: LA Metro Bike Share

Data Source: `https://bikeshare.metro.net/about/data/`

Metro Bike Share is a bicycle sharing system in the city of Los Angeles, California Metropolitan area. The bike sharing system was launched in 2016. It is administered by the Los Angeles County Metropolitan Transportation Authority and is operated by Bicycle Transit Systems.

The reason this topic is selected is because there are some eminent positive positive effects of bike sharing related to environmental sustainability and the personal health benefits experienced by users. Bike share schemes like Metro (LA), Pronto (Seattle) etc. are intended to âĂIJgreeningâĂİ of the city; ideally, more bikes can equal fewer cars, which leads to a reduction in greenhouse gases and an improvement in local air quality. It also has significant health benefits.

2 PROBLEM STATEMENT

Using the LA Bike Share data try to predict âĂIJPricing plan maximizing total revenue from bike usageâĂİ based on tickets and pass-types, trip duration, and number of trips by region. If possible, create a mobility modelling based on the historical bike usage data.

3 PROJECT SCOPE

Scope of the project is limited to data discovery, data acquisition, data quality, data preparation, data assumptions, exploratory data analysis, predictive model creation, visual representation and final project report creation.

TECHNICAL APPROACH

4 ANALYSIS

Analysis will consist of data discovery and exploration. Specifically,

- Finding the right data source

- Creating project structure
- Data analysis, munging, transformation
- Exploratory data analysis
- Predictive Analytics

5 Requirement Development

Business and Functional Requirement: Metro Bike Share company needs an effective way to stuff its bike and roll out new membership or continue existing membership so that their revenue is maximized and high level of customer satisfaction is achieved through membership offerings and bike availability.

Technical Requirement: Three basic data sets are required for the model. They are:

1. Bike Share trip data with duration, geocoding, and membership type.
2. Station details with code and location.
3. Weather data for the period in analysis.

6 Model Deployment

The scope of this project is to build and evaluate the model. Deployment is out of scope, since we do not have a cloud solution for deployment.

7 Testing and Evaluation

The model should be tested and evaluated based on the test and train data randomly sampled from the project data set.

8 Specific Exclusions from Scope

Depending on the processing performance, we may or may not include all trip data since 2016.

Expected result

Once finished, the project will tell us which membership type has potential to earn the most revenue for the bike share company, as well as which route can have the maximum traffic. This will help the bike share company to maximize its revenue, and optimize its bike stuffing in different stations.

MANAGEMENT APPROACH

9 PROJECT PLAN

The project will be released in an incremental MVP (minimal viable product) approach as the project is phased as assignments over few weeks. The weekly assignment also governs what will be released each week.

10 PROJECT RISK

Since we would be working on this project along with other major assignments and the summer term is shorter, time-line is one risk that we need to mitigate. Collaboration from different geographies pertains the risk of not having effective communication and collaboration. For the project, would be collecting open data from internet, hence data quality might be a risk. During the analysis, if we need any supplemental data, finding that data might be a risk, given there is no central data source.

APPROVAL AND AUTHORITY TO PROCEED

We approve the project as described above, and authorize the team to proceed.

Name	Title	Date
Anirban Pal	Member	6/11/2019
Gangadhar Dhulipala	Member	6/16/2019
Raghu Raman Nanduri	Member	6/16/2019