

Business Buddy - Tool to help businesses make informed decisions

Sunday, December 15, 2019 1:40 PM

Name : Dana Moukheiber (danamouk@buffalo.edu), Nikita Goswami(nikitagow@buffalo.edu), Saurabh Mahindre(smahindr@buffalo.edu)

Course : Programming and Database Fundamentals for Data Scientists (EAS503)

Title : Analysis of Customer Reviews and User Data

Aims & Objectives of the Project

Our objective is to generate actionable insights using huge unstructured data like reviews posted by users to help existing and future business owners. We have used various data analysis techniques to extract sentiments from positive and negative reviews and combined it with businesses and user profile data to give a holistic view of how well a business is doing in various demographics. This will empower the stakeholders including businesses to drive products, services, customer pain points, market opportunity related decisions backed by data. Our analysis will also help generate insights to make recommendations for prospective customers.

Data: For analysis, we have taken data from the application "Yelp", The data is specific to US and Canada market. The data includes Reviews given by Users, Check-in Data of users for a business and metadata of the businesses.

Our analysis will help the below stakeholders in the following ways :

1. Businesses
 - Business get to know areas for improvement by getting to know customer pain points.
 - Businesses can find out avenue to increase customer base and profit based on timing and day of the week.
 - Gauge how well competitors nearby are doing.
2. Customers/ Tourists
 - Best and closest nearby locations to visit in Canada and US
3. Investors
 - Gauge idea of scalability of the market and business ratings over the years from time series analysis .

Details of Analysis

1. We have inserted the data into the database : The database has the following tables - Business, Reviews and Checkin
2. We have analyzed which cities have the most well reviewed business and what kind of business work well in all major cities based on customer's reviews and average ratings of the place
3. We have performed Natural language processing on the good and bad reviews of the business and highlighted the positives and negatives of the business using Word Clouds
4. Analyzed change in ratings over time for top and bottom performing businesses.
5. Analyzed user check-ins for the businesses to look for patterns over different days of the week and different hours in the day. This would help the business think of offers to be given during the off days and hours to increase their business in the dull hours.
6. Analyzed sentiment scores for each review to help us understand if they stand close to the overall average review of a business.

Techniques used for analysis:

1. Analysis : Time Series Analysis, Sentiment Analysis using Affin Scores, Word Clouds, Topic Modelling using Latent Dirichlet Allocation
2. Visualizations : Matplotlib Histogram, Plotly Pi charts, Seaborn Bar charts, Basemap, Word Cloud

Conclusions

1. Las Vegas has the most number of businesses mentioned on the Yelp Platform from where we took our data. Phoenix and Canada are following closely. We found out what kind of businesses work well in these top cities
2. On studying user reviews we found out that most users prefer to give above average ratings mostly between 3.5 and 4 (on a scale of 5)
3. We analyzed the topmost reviewed business called "Earl of Sandwich" located in Las Vegas. This place had maximum good ratings and reviews. For this place , we found out that :

- a. Customers like the quality of the food and the favorites from this place are some sauces. People also like the price of food at this place.
- b. The negative things customers said about this place were related to the time they had to spend to give their orders because of a queue and customers do not seem to like the service at this place so much.
- c. On segregating the reviews into 8 categories, we got to know about what people are specifically talking about various things: They seem to like the tuna fish and sushi here, Chicken Dishes and soup are being talked about and the location, logistics, ordering are the other topics which have reviews tagged to them. This analysis helped us zero down to what customers feel is good and not so good at this place.
- d. On analyzing the reviews over the years, we realized that in a particular month each year the reviews of the place are going down. This can be because of some tourist season due to which the waiting time is too long and the quality of food is deteriorating.
- e. The place is packed during the evening hours on weekdays and during morning and evening hours on weekends. They can have some offers running in their dull hours to attract more customers.
4. For the worst rated business "Fox Rent a Car", we did similar analysis as above and found out that this place had better reviews until 2016 but since then it has deteriorated and customers have a lot of complaints with their services.

Future Enhancements Planned

1. We would include more user data for the analysis and incorporate collaborative filtering capabilities to help customers find the best business
2. Use Word2Vec and Deep Learning for providing suggestions and analysis