Capstone Project Submission

Team Member's Name, Email and Contribution:

1. Sharad Shantaram Tawade
Email:- tawadesharad13@gmail.com
Mob:- 7507121013

Contribution:-
1. Data Wrangling.
2. Exploring And Visualizing the data.
3. Text-Preprocessing- Remove punctuation, Remove stopwords, Lemmatization
4. Sentiment Analysis
    - MultinomialNB
    - Random Forest
    - XGBoost
    - SVM(Support Vector Machine)
5. LDA (Latent Dirichlet Allocation)
6. Clustering
    - K-means Clustering
    - K-Means with PCA
7. Conclusion
8. Documentation

2. Sagar Malik
Email:-sagarmalik2582@gmail.com
Mob:- 9560203146

Contribution:-
1. Data Wrangling.
2. Exploring And Visualizing the data.
3. Text pre-processing
4. Sentiment Analysis
5. KMeans
6. LDA
7. XGBoost
8. Support Vector Classifier (SVC)
9. Random Forest

3. Vinay Kumar
Email:-vinaykumar.nigudgi@gmail.com
Mob:- 8618303442

Contribution:-
1. Data Wrangling.
2. Exploring And Visualizing the data.
3. Text pre-processing
4. Sentiment Analysis
5. KMeans
6. LDA
7. XGBoost
8. Support Vector Classifier (SVC)

Github Link:- Link to GitHub repo

Summary

In today's digital era where everybody has their own opinion and so can post various feedbacks regarding different products. Reviews are critical in nature and can affect the market value of a product. Amazon, Zomato, Quora, Twitter, etc . are few such platforms that provide space for product reviews. Since reviews are textual and diversified; a collective representation will help users to summarize their opinions. Honest food reviews and their analysis is still a challenge.

In this Project we focus on Customers and Company, we analyze the sentiments of the reviews given by the customer in the data and make some useful conclusions in the form of Visualization. Also, cluster the zomato restaurants into different segments using clustering algorithms.

The main objective is to cluster the restaurants into segments and sentiment analysis of reviews. Also, the data has valuable information around cuisine and costs which can be used in cost vs. benefit analysis.

We here use many classifiers for sentiment analysis like MultinomialNB, Random Forest, XGB and SVM. We use clustering techniques K-Means and PCA (Principal Components Analysis) for cuisines.