Sentiment Analysis & Rating Prediction for Customer Reviews

Abstract

Problem statement:

The Businesses get a lot of reviews in form of comments and stars sometimes the websites do not have a rating system but rather just a comment system. So going through all the comments is not efficient. Our project seeks to provide solution to these businesses and predict the ratings along with the sentiments of customers from the comment.

Literature Survey:

Sunmin Lee et al [1] The study focused on the initial stage of the algorithm by answering the research question that can the Bidirectional Encoder Representations from Transformers model determine whether a customer's review on Yelp is positive or negative, and the degree of said positivity or negativity, based on the review's content. Boya Yu et al [2] The main approach used in this paper is to use a support vector machine (SVM) model to decipher the sentiment tendency of each review from word frequency. Word scores generated from the SVM models are further processed into a polarity index indicating the significance of each word for special types of restaurants. Customers overall tend to express more sentiment regarding service. Anish A. Parikh et al [3] This study's purpose was to identify factors of usage, trust, influence, and contribution of restaurant reviews on Yelp.com. This study found that information search reduction and community membership were the greatest factors encouraging Yelp.com use.

Dataset: Yelp Dataset

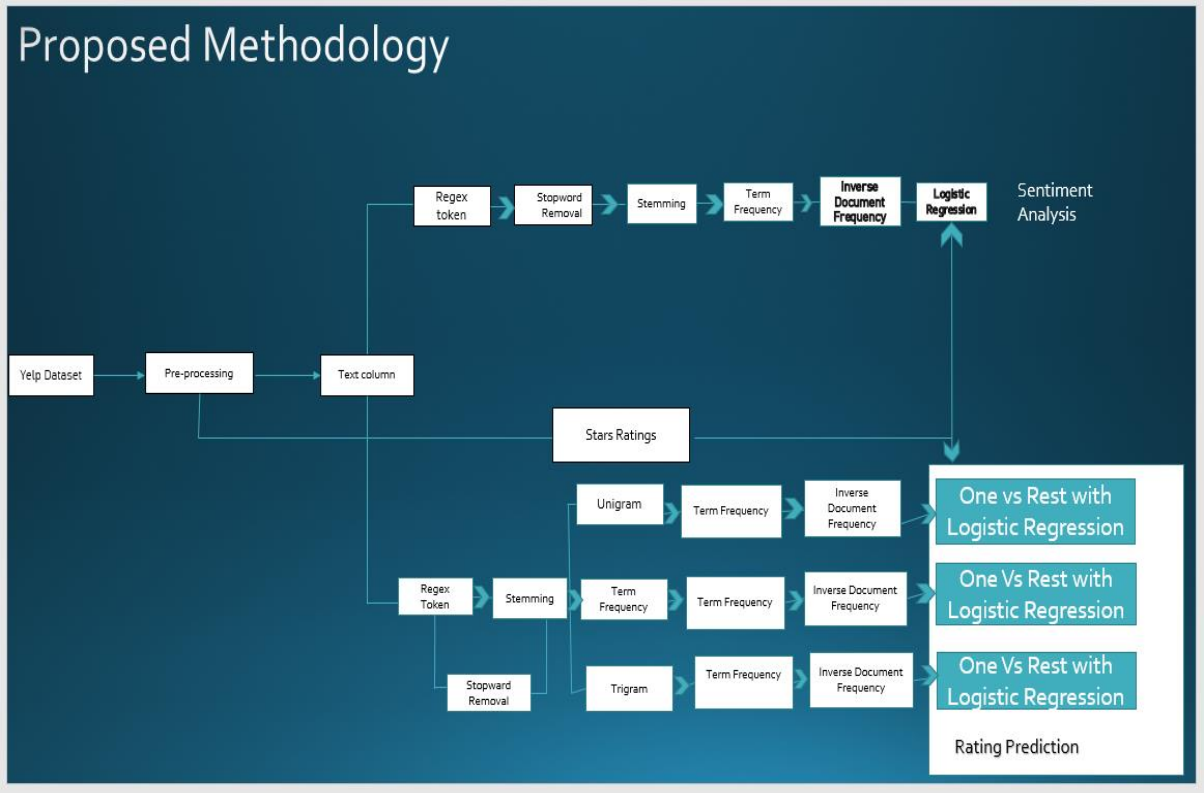
This dataset is a subset of Yelp's businesses, reviews, and user data. It was originally put together for the Yelp Dataset Challenge which is a chance for students to conduct research or analysis on Yelp's data and share their discoveries. In the most recent dataset, you'll find information about businesses across 8 metropolitan areas in the USA and Canada.

Methodology:

In this experiment we are using Natural Language Processing: Natural language processing (NLP) refers to the branch of computer science—and more specifically, the branch of artificial intelligence or AI—concerned with giving computers the ability to understand text and spoken words in much the same way human beings can.

NLP combines computational linguistics—rule-based modelling of human language—with statistical, machine learning, and deep learning models. Together, these technologies enable computers to process human language in the form of text or voice data and to 'understand' its full meaning, complete with the speaker or writer's intent and sentiment.

The proposed methodology for the experiment is as shown in the flow chart:



References:

1. Sunmin Lee "Sentiment Analysis Using BERT on Yelp Restaurant Reviews" Department of Computer and Information Technology West Lafayette, Indiana August 2022.
2. Boya Yu, Jiaxu Zhou, Yi Zhang, Yunong Cao "Identifying Restaurant Features via Sentiment Analysis on Yelp Reviews" Center for Urban Science & Progress New York University, New York, The United States.
3. Anish A. Parikh,[1] Carl Behnke,[2] Doug Nelson,[2] Mihaela Vorvoreanu,[3] and Barbara Almanza[2] "A Qualitative Assessment of Yelp.Com Users' Motivations to Submit and Read Restaurant Reviews" [1]Department of Management, Montclair State University, Montclair, New Jersey, USA [2]School of Hospitality and Tourism Management, Purdue University, West Lafayette,Indiana,USA [3]Department of Communication, Purdue University, West Lafayette, Indiana, USA.

Group Members:

1. KADAM VINOD MARUTI (220960925016)
2. KENGE DURGESH RAOSAHEB (220960925017)
3. KOTHEKAR HARSHADA VINAYAK (220960925019)
4. KULKARNI SHREYA AVINASH (220960925020)
5. KUNAL BENDALE (220960925021)
6. PATEL BHAUMIK NARENDRABHAI (220960925024)

Group No: 04