Customer Reviews of Amazon Products

Ayushi Agarwal
Computer Science Department
PES University
Jaipur,India
ayushiawesome10@gmail.com

Harsha Bothra
Computer Science Department
PES University
Kolkata,India
harshabothra99@gmail.com

Vismitha Hiremath
Computer Science Department
PES University
Bangalore,India
vismithahiremath@gmail.com

Abstract—Online E-commerce websites like Amazon, Filpkart uses different recommendation models to provide different suggestions to different users. Amazon currently uses item-to-item collaborative filtering, which scales to massive data sets and produces high-quality recommendations in real time. This type of filtering matches each of the user's purchased and rated items to similar items, then combines those similar items into a recommendation list for the user. In this project we are going to build recommendation model for the electronics products of Amazon.

I. Introduction

Our project Amazon reviews on kindle dataset use customer feedback reviews or ratings on the product, where we utilize techniques such as opinion mining, text mining and sentiments, which has affected the surrounded world by changing their opinion on a specific product. Sentiment analysis is a series of methods, techniques, and tools about detecting and extracting subjective information from the user/customer, such as opinion and attitudes, from language, helping in finding the mood of the customers about a purchasing of a particular product or topic in relation to entities such as products, services, organizations, events, topics and their different attributes.

II. Ease of Use

Our aim in the project was to draw sentiment analysis by taking feedback from the customers on amazon kindle reviews dataset and infer necessary information to extract the requirements of the user to best fit their needs. In today's world with increasing advance technology and never-ending demand of the users, it becomes extremely essential to built and predict essential models that can help curb the requirements of the user and accommodate with the best-suited services. We built our model to increase the efficiency and accuracy to increase the returns for the business and smoother experience.

A. Consumers want to find useful reviews as quickly as possible by the help and assistance of the rating system. Therefore, models that are capable to predict the user ratings from the text review are critically important. Getting an overall sense of a textual review could in turn improve the consumer experience. Also, it can help businesses to increase sales, and improve the product by understanding customer's needs.

III. The dataset had predecessor work which included cleaning the dataset (to reduce the missing and the null values), analysis of data, Transforming data Time (i.e parse review data to, Parse reviewDataAddedto, Parse ReviewDataSeen to),likert scale Analysis(i.e Detractors,Passive,Promoter),we also found the NPS net promoter score of Amazon.Different visualizations carried out to interpret the data for fruitful results.

IV. Prepare Your Paper Before Styling

We examined if all ratings were given by the same customer or were there any inconsistency between them. This was clarified after visualizing and juxtaposing with the help of bar graph and pie chart.

We had an appertain stemmed algorithm because, lower attributes means more information density in the compressed columns. Consequently it was found that the stemmed algorithm has attributes that were 18% lower than that of lemmatized approach.

The whole process includes summarization in three steps which are as follows:

(1) Product feature-based, which is given by the customer.

(2) Identify expected features in each opinion sentence or review

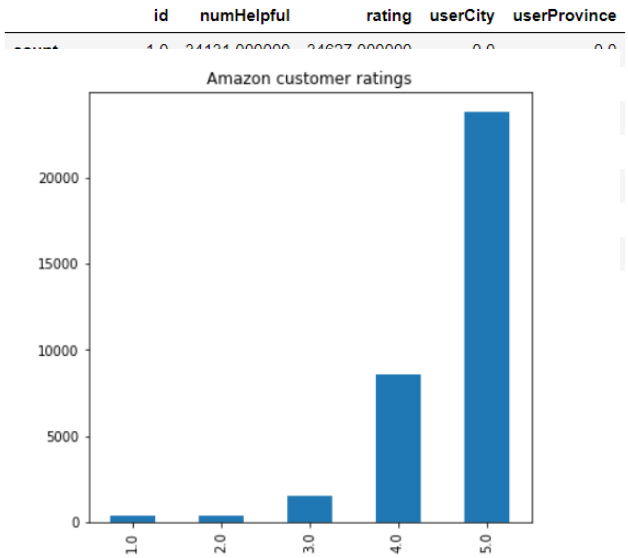
(3) Finding out whether the feature/opinion is positive, negative, or neutral and finally, the summary will be created.

A. Abbreviations and Acronyms

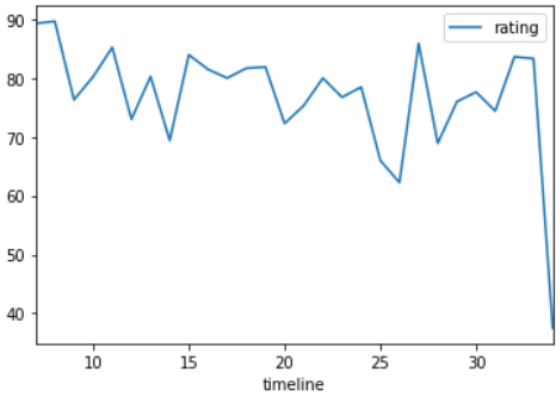
This project has significant solutions to customer reviews. We started the project with Understanding and cleaning the data.we checked for null values and dropped columns which aren't useful.the next phase was to theorize whether rating are genuine, which included to check if one user is trying to give all rating , accordingly will the distribution look for bulk users and the number of users in bulk.In the next stage we found the NPS net promoter score of

amazon. We figured the NPS score and calculated it for amazon. We picked one variation product in which we drilled and inspected its property.

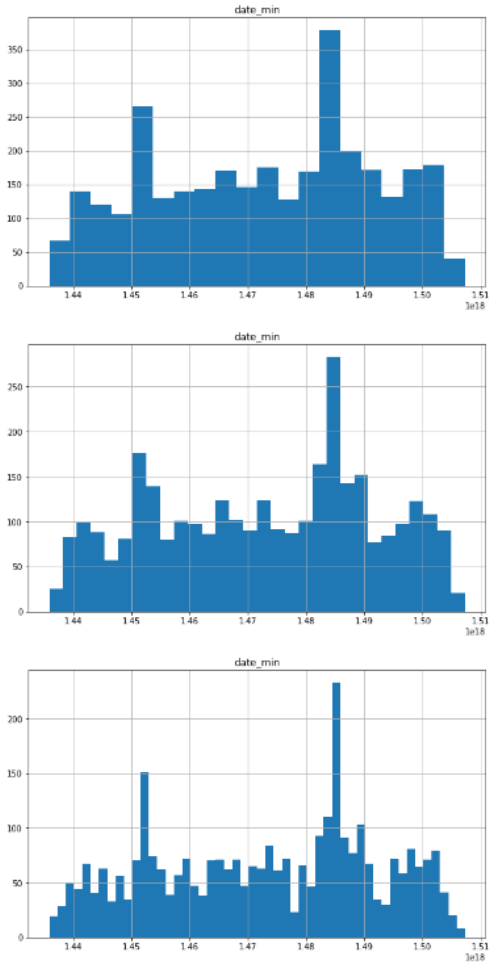
We plotted time series for review, which comprised handling date time text, plot time series on a graph and graph with various internals. After this analysis we had an insight that “January month has the highest number of peaks >> Activity is high >> More Sales during Jan as we had high degree of variance in reviews added over time”. the following step was to predict recommendations based on reviews content. For this step we had to make clean functions (i.e remove punctuations, remove stopwords,stem vs Lemmatize),create a TFIDF vectorizer, Features, Understand and explore sentiment analysis was an important slice. After the Random Forest classifier we could terminate by checking the score and analysing the data set with all the predictions.



The equations are an exception to the prescribed specifications of this template. You will need to determine Reviews not only have the power to influence consumer decisions but can strengthen a



company's credibility. Reviews have the power to gain customer trust, and they encourage people to interact with the company. Customer interaction ultimately leads to improved profits for businesses. Customer satisfaction survey is a tool that enables companies to gain insight of customers needs and requirements, their loyalty, as well as obtain useful



feedback, whether in a positive or in a negative light, its benefits are vastly employed in modern business to enable companies to achieve more customer

Buying Decisions Are Based On Customer Reviews. Positive or negative reviews can significantly impact sales because consumers will look for feedback before making a purchase decision. If they find the reviews reliable and accurate, they will most likely listen to them. The model turned to give high accuracy and precision of (*% each with a recall of 99%.Knowing and understanding customer needs is at the centre of every successful business, whether it sells directly to individuals or other businesses. Once you have this knowledge, you can use it to persuade potential and existing customers that buying from you is in their best interests.

Portrayal and visualizing insights from the dataset on various framework and constrast the results obtained earlier and enhance it was the pivot ,we tried to achieve in the project. Better plots, graph and charts helped to adopt strategies to increase the performance of the model and earn better results on future data also. We also subsume technique/task like lemmatization and stemmed process, calculate the NPS score and carried out an inspection on time series data to check and verify the dominion of our dataset in line with measure.

	id	numHelpful	rating	userCity	userProvince
count	1.0	34131.000000	34627.000000	0.0	0.0
mean	111372787.0	0.630248	4.584573	NaN	NaN
std	NaN	13.215775	0.735653	NaN	NaN
min	111372787.0	0.000000	1.000000	NaN	NaN
25%	111372787.0	0.000000	4.000000	NaN	NaN
50%	111372787.0	0.000000	5.000000	NaN	NaN
75%	111372787.0	0.000000	5.000000	NaN	NaN
max	111372787.0	814.000000	5.000000	NaN	NaN