Feedback Analyzer

Aditya Vardhan **Ikya Venus** **Kethavath Naveen** **NKVS Kiran**
Indian Institute of Technology Guwahati
Group No:
{adit170101003, ikya170101024, navee170101032, sai170101039}@iitg.ac.in

Abstract

With growing technology all around us one of the key aspect emerged is to manage the data. Enterprises nowadays are keen in getting customers intentions through their feedback. Earlier means of receiving feedback from customers and drawing conclusion over millions of feedback is a lump sum process. So there emerged a need to automate the process of acquiring intention of the customer behind the feedback to know the value of ones product. Our idea is to design such a model which analyze the intention of the customer behind the feed back and deliver automated reply to the customer in real time. Sentiment analysis is part of the Natural Language Processing (NLP) techniques that consists in extracting emotions related to some raw texts. This is usually used on social media posts and customer reviews in order to automatically understand if some users are positive or negative and why.

1 Introduction

Feedback is one of the prominent things in marketing by which we can understand the brand's product value in the market. We want to use NLP ideas like sentiment analysis to automate the understanding of customers intentions, the key aspects of a product that customers care for and their underlying intentions and reactions behind it. Our model is a prototype of the idea which leverages NLP techniques to understand the intentions behind the feedback received. Word embeddings are effective intermediate representations for capturing semantic regularities between words in natural language processing (NLP) tasks. We propose sentiment-aware word embedding for emotional classification, which consists of integrating sentiment evidence within the emotional embedding component of a term vector. We take advantage of the multiple types of emotional knowledge, just as the existing emotional lexicon, to build emotional word vectors to represent emotional information.

2 Data

The base of generating any model in the machine learning is the data. Data drives a model to be good or bad. So our key step is collect data to feed our model. So we need a good amount of data to feed our model. So we are using Amazon product reviews API to get the product reviews and use the rating to label it as good or bad one. We are considering ¡3 as negative review and remaining as positive reviews.

We split our data into two parts one for training and one for testing purpose. For training purpose we used the rating to classify it as a negative or positive feedback.

3 Progress:

3.1.Data Collection:

We made use of amazon product reviews API to use as corpus for our model. We split the data into two parts one for training and one for testing.

3.2.Classifying data:

To classify data we used the idea that ¡3 star reviews are negative feedback and remaining are positive feedback.

3.3.Developing model: We used star rating to classify data as -ve and +ve reviews

We started by adding sentiment analysis features because we can guess that customers reviews are highly linked to how they felt about the product on amazon. We use Vader, which is a part of the NLTK module designed for sentiment analysis. Vader uses a lexicon of words to find which ones are positives or negatives. It also takes into account the context of the sentences to determine the sentiment scores. For each text, Vader returns 4 values: A neutrality score, positivity score, negativity score and an overall score that summarizes the previous scores.

4 Implementation

Process followed so far for sentiment analysis: Step 1 : Collecting Data/feedback Step 2 : Tokenizing the data (No of characters in the text / No of words in the text) Step 3 : Removing the unnecessary tokens(eg.stop words) Step 4 : Classifying each word as +ve or -ve(Using specific terms) Step 5 : Trianing the data with supervised learning model Step 6 : Calculating the +ve or -ve percentage.

The next step consist in extracting vector representations for every review. The module Gensim creates a numerical vector representation of every word in the corpus by using the contexts in which they appear (Word2Vec). This is performed using shallow neural networks. What's interesting is that similar words will have similar representation vectors. Each text can also be transformed into numerical vectors using the word vectors (Doc2Vec). Same texts will also have similar representations and that is why we can use those vectors as training features. We first have to train a Doc2Vec model by feeding in our text data. By applying this model on our reviews, we can get those representation vectors. The next step consist in extracting vector representations for every review. The module Gensim creates a numerical vector representation of every word in the corpus by using the contexts in which they appear (Word2Vec). This is performed using shallow neural networks. What's interesting is that similar words will have similar representation vectors. Each text can also be transformed into numerical vectors using the word vectors (Doc2Vec). Same texts will also have similar representations and that is why we can use those vectors as training features. We first have to train a Doc2Vec model by feeding in our text data. By applying this model on our reviews, we can get those representation vectors.

5 Expecting to do

Organising Feedback into genres: Along with this we will implement, All feedbacks using semantics in the feedback, we are organising them into different types of logical structures and genres example Emotional feedback, Happy feedback, Angry feedback etc. Add auto reply feature: With data classified into different genres using logical structures, we will implement auto reply feature with a fixed data of replies in the dataset. Extracting suggestion using key aspects of product.

6 Analysing Amazon Product Reviews Data

Why we need this?? E-commerce has revolutionized the way we shop. Everything we search are just a few clicks away. Items are being delivered within a matter of days (sometimes even the same day!). Amazon is one such platform that has a huge impact in E-commerce industry.
Many times while purchazing an item online, people see the product's, ratings reviews before deciding whether to buy it or not.
Online product reviews are a great source of information for the Organization. They can analyze this data to imporve the customer satisfaction and increase revenue.
Amazon Customer Reviews is one of Amazons iconic products. In a period of over two decades , millions of Amazon customers have contributed over a hundred million reviews to express opinions and describe their experiences regarding products on the Amazon.com website. Over 130 million customer reviews are available to researchers as part of this dataset.// With such Humongous review data available, My target was to analyze the data, find patterns to classify the products,find the best product and recommend it to the user. For that I have used below approaches. **1.Sentiment Analysis :** To certify whether the product is be good or bad. **2.Recommender System :** Recommend the next best product to the user.

Importance of Online Reviews : Many times while purchasing an item online people see the ratings, but ratings alone do not give a complete picture of the products we wish to purchase. So, as a precautionary measure, people tend to read a product's reviews before deciding whether to buy it or not. Online product reviews are a great source of information for consumers. From the seller's point of view, online reviews can be used to gauge the consumers' feedback on the products or services they are selling.

Problems Of Fake Reviews : Online reviews have become an important factor when people make purchase and business decisions. The increasing popularity of online reviews also stimulates the business of fake review writing, which refers to paid human writers producing deceptive reviews to influence readers' opinions. Such fake reviews can create problems to customers who are accustomed to reading reviews before making a final purchase decision as the decisions are possibly influenced by non-consumers. This project aims to eliminate fake reviews as accurate as possible.

Sentiment Analysis : Customer satisfaction is the main key to success in business. its key for companies to pay close attention to Voice of Customer to improve the customer experience. By analyzing and getting insights from customer feedback, companies have better information to make strategic decisions, an accurate understanding of what the customer actually wants and, as a result, a better experience for the customer. To achieve this we can use the concept of sentiment analysis on the product reviews that we get as a feedback from customer. Sentiment Analysis is the automated process of understanding the sentiment or opinion of a given text. This machine learning tool can provide insights by automatically analyzing product reviews and separating them into tags: positive, neutral, negative.

Recommender System : Online stores have millions of products available in their catalogs. Finding the right product becomes difficult because of this Information overload. Users get confused and this puts a cognitive overload on the user in choosing a product. Recommender systems help customers by suggesting probable list of products from which they can easily select the right one. They make customers aware of new and/or similar products available for purchase by providing comparable costs, features, delivery times etc.
In many ways, recommender systems were a catalyst for the current popularity of machine learning. One of Amazon's earliest successes was the Customers who bought this, also bought feature, while the million dollar Netflix Prize spurred research, raised public awareness, and inspired numerous other data science competitions.

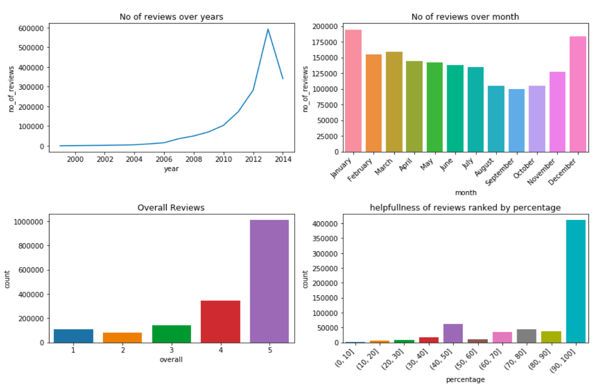
Dataset : The dataset contains the customer review text with accompanying metadata, consisting of three major components:
Amazon Customer Reviews (a.k.a. Product Reviews) is a database with data from a a period of over two decades from 1995 to 2014, millions of Amazon customers have contributed over a hundred million reviews to express opinions and describe their experiences regarding products on the Amazon.com website. Over 130+ million customer reviews are available to researchers as part of this dataset.
As this is a very huge dataset(140GB), I have chosen 'Electronics' category to perform Analysis and generate insights that address the problem statement.

Proposed Flow of The Model
a.Data Cleaning
b.Data Preprocessing
c.Analyzing Reviews by Exploratory Data Analysis
d.Sentiment Analysis of Reviews
e.Product Recommendations based on the reviews and overall score.
Building a model that tries to solve below problems

1.Sentiment Analysis on Product reviews.
2.Product Recommendations based on the reviews and overall score.

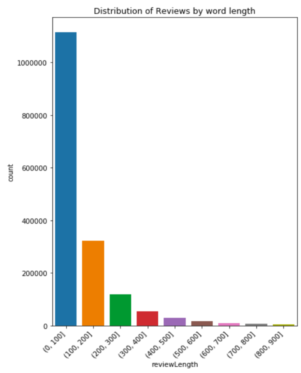
Exploratory Data Analysis

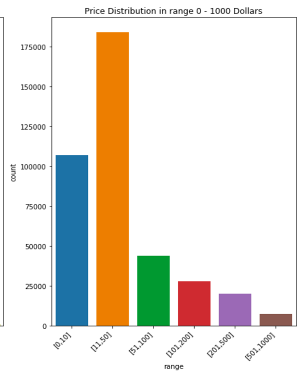
Rating Trend over the years : There is an increasing trend for number of ratings given by the users to products on amazon which indicates that a greater number of users started using the amazon e-commerce site for online shopping and a greater number of users started giving feedback on the products purchased. there is a increase in number of ratings given by users from 2012 to 2014. Notice that the peak on 2013. further analysis will show that this happened on November of that year. two major events support these findings. the first one is that on that month, Amazon began to offer Sunday delivery option for purchases.That surely resulted in lots of new members and new ratings reviews. The second one is obviously black Friday.
Observation from Exploratory Data Anlysis :-



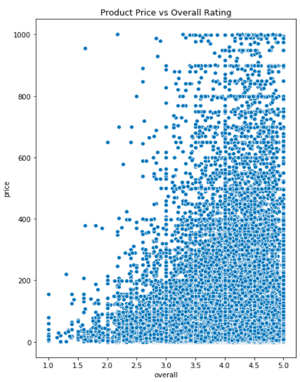
- we can see the reviews trend over the years which shows the increased usage of website which inturn tells us the brand Amazon is now.
- The number of reviews generated per month in every year, gives us a basic idea of online shopping pattern.
- Analysing the overall rating of the products, to get most rated product, most popular brand etc.
- see how helpful the reviews are from their helpfullness score.

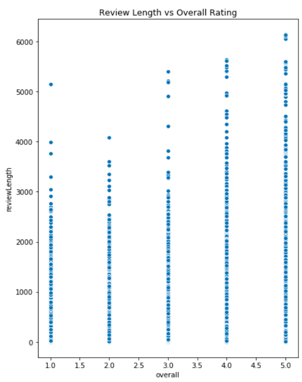
Distribution of Overall ratings : Many users have given a rating of 5 to products followed by 4 and 3 whereas very few users have given a low rating of 1 or 2. For this look at below plot,



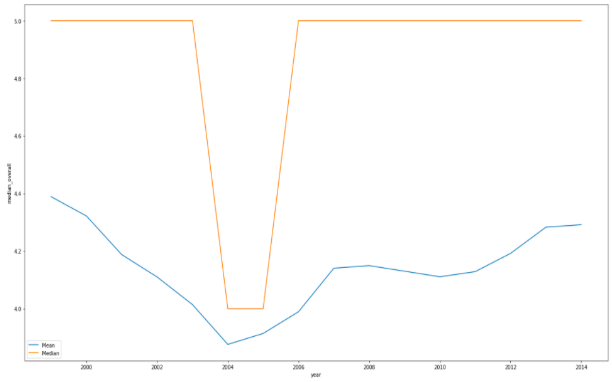


Most of the reviews come in the range of 3-5 for a price range of 200-400. And rating and review length are related to each other. For this look at below plot,,

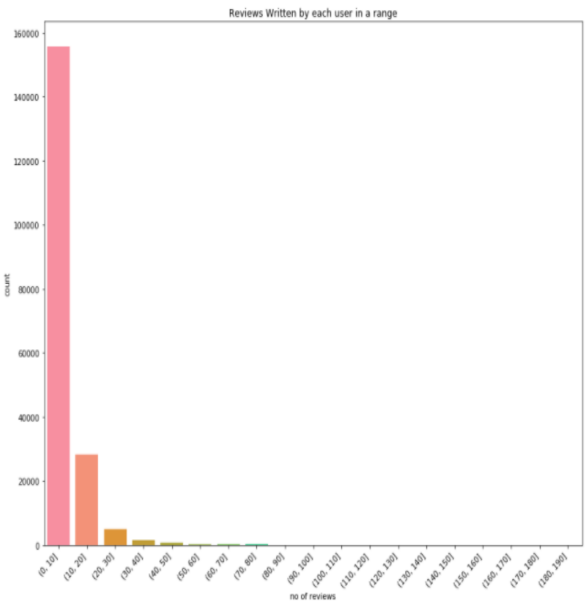




we can infer that over the years 2000 to 2004, the mean rating of the products has reduced and then increased at a slow rate till 2014 but still is much lower than before 2000. Median of ratings given to products remains at 5 from 2000 to 2014 except for years 2004 and 2006. For this look at below plot,



Reviews Written by each user in a range : On an average each user gives 5 reviews and most reviewed user has given 430 reviews for all the different products, facts and figures are shown below in recommender system as it gives us a better understanding of the user.



Data Cleaning

Data preprocessing and cleaning is an important step before any text processing task, in this step, we will remove the punctuations, stopwords and normalize the reviews as much as possible.

Handling Duplicate Data : There exist a lot of duplicates wherein the different products is reviewed by same user at the same time The product ID may be different but the product is similar with different variant.

Text Preprocessing : As the review is mostly text data, we might need to clean the data to gain some useful insights from the data.

The various text preprocessing steps are:
1.Tokenization
2.Lower casing
3.Stop words removal
4.Stemming
5.Lemmatization
These various text preprocessing steps are widely used for dimensionality reduction.

Word Cloud : Word Clouds are the graphical representation of word frequency that give greater prominence to words that appear more frequently in a source text. the large the word in the visual the more common the word was in the document.


```
Word Cloud processing completed in
CPU times: user 324 ms, sys: 9.65 ms, total: 334 ms
Wall time: 341 ms
```

Sentiment Analysis

Sentiment Analysis is the automated process of understanding the sentiment or opinion of a given text. This machine learning tool can provide insights by automatically analyzing product reviews and separating them into tags: Positive, Neutral, Negative.

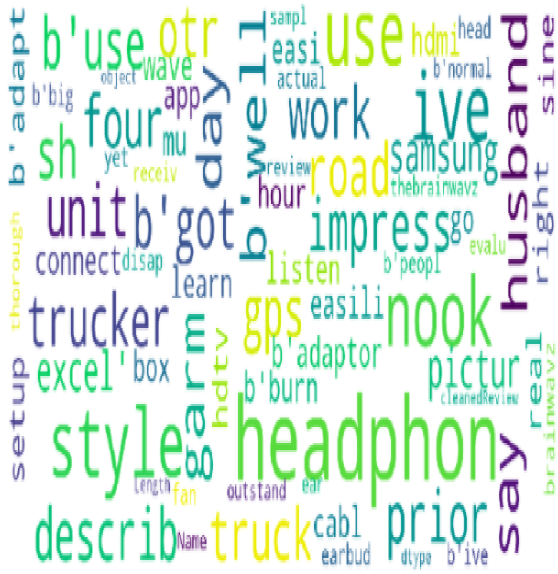
Classifying reviews based on sentiment using VaderSentiment : In this part, I have used a prebuilt library VaderSentiment which is used in predicting the sentiment of a review based on the lexicon arrangement of the words in a review. Also different statergies to actually predict the sentiment for each review by classifying it into Positive, Neutral and Negative reviews.

- Using naive bayes classifier to generate probabilities for each document
- Based on generated sentiment score on each document, we can classify the review into 3 categories(positive negative and neutral)

As we have generated sentiment scores for all the reviews we can now see what are the most repeated words for positive and negative reviews using a word cloud.

Positive Reviews

```
#---- Positive Reviews
%%time
show_wordcloud(positive_data['cleanedReview'])
print()
print("Word Cloud processing completed in ")
```




```
Word Cloud processing completed in
CPU times: user 320 ms, sys: 11.5 ms, total: 332 ms
Wall time: 331 ms
```

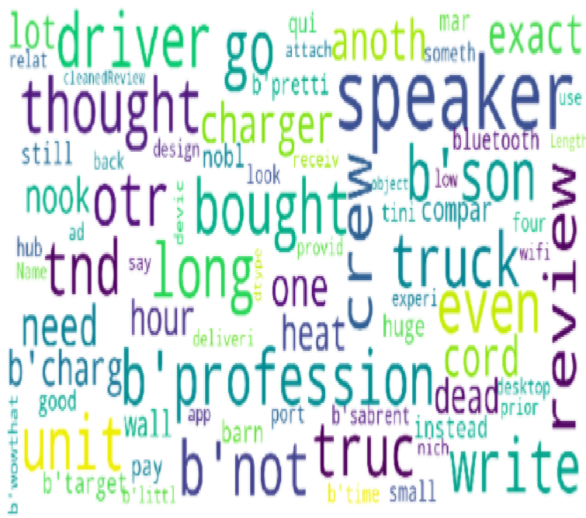
Negative Reviews

```
#---- Negative Reviews
%%time
show_wordcloud(negative_data['cleanedReview'])
print()
print("Word Cloud processing completed in ")
```


```
Word Cloud processing completed in
CPU times: user 332 ms, sys: 7.67 ms, total: 340 ms
Wall time: 339 ms
```

Neutral Reviews

```
#---- Neutral Reviews
%%time
show_wordcloud(neutral_data['cleanedReview'])
print()
print("Word Cloud processing completed in ")
```




```
Word Cloud processing completed in
CPU times: user 318 ms, sys: 11.5 ms, total: 330 ms
Wall time: 330 ms
```

Assign new dimension to each word and give the word counts

TfidfTransformer().fit transform to fit the train and test data

We use below methods to classify the sentiment of the data.

- Multinomial Naïve Bayes learning method
- Logistic regression learning method

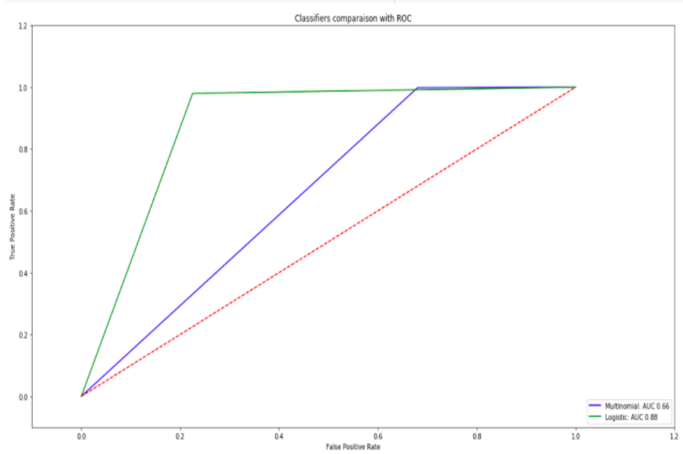
1.Naive Bayes classifiers are a family of simple ”probabilistic classifiers” based on applying Bayes' theorem with strong (naïve) independence assumptions between the features.

2.Logistic regression is a statistical model that in its basic form uses a logistic function to model a binary dependent variable, although many more complex extensions exist. In regression analysis, logistic regression (or logit regression) is estimating the parameters of a logistic model (a form of binary regression).

Evaluating The Results

- using ROC curve

Using the curve we can see which model has better performance and use that model for testing purpose.

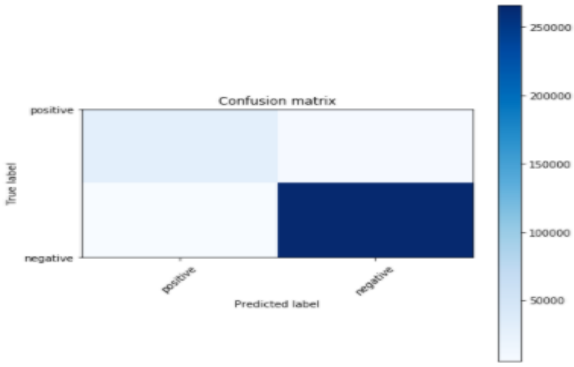


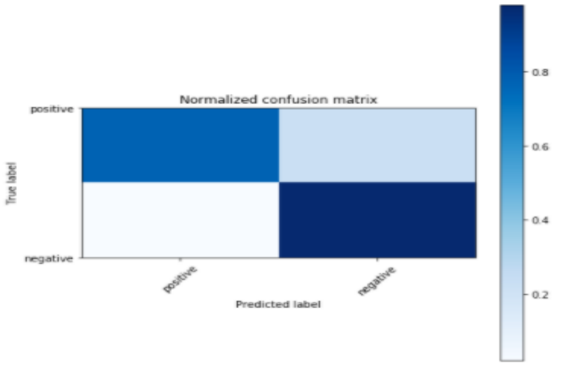

```
              precision    recall  f1-score   support

    positive       0.84      0.77      0.81     38049
    negative       0.97      0.98      0.97    271338

    accuracy                           0.95    309387
   macro avg       0.91      0.88      0.89    309387
weighted avg       0.95      0.95      0.95    309387
```

Test Accuracy : 0.95440981533257

we can see in the above figure that the logistic regression clearly outperforms other model. so lets see the accuracy of logistic regression.
Visualize the accuracy, recall and f1-score for Logistic Regression.

Plotting the Confusion Matrix



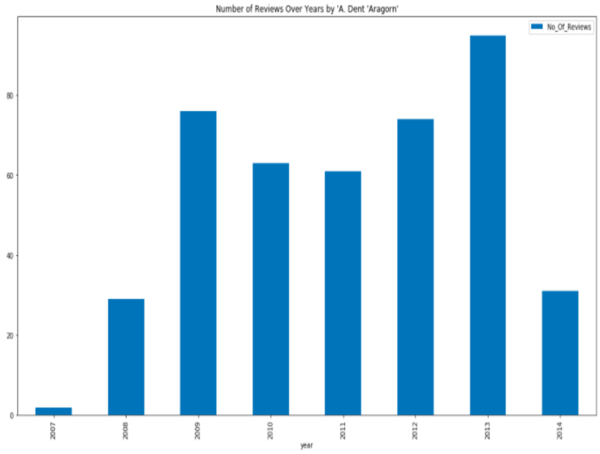


Getting the words that classify the best and worst features. We us the logistic model as it gives the best results.

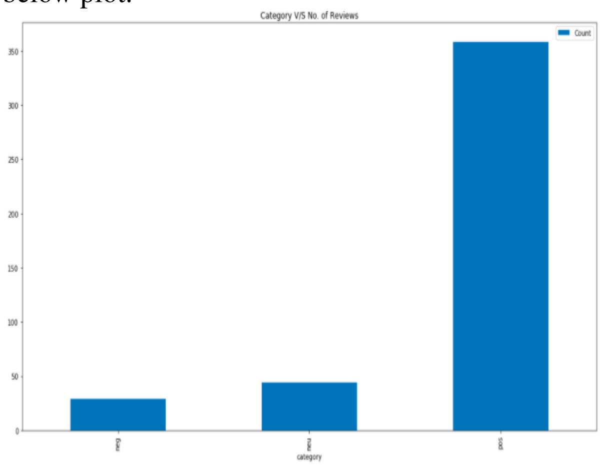
Sentiment analysis is performed using logistic regression and Naive Bayes. Also, the user behavior is analyzed and the popular words used by the users are determined.

Recommendation System : Recommender systems help customers by suggesting probable list of products from which they can easily select the right one. They make customers aware of new and/or similar products available for purchase by providing comparable costs, features, delivery times etc.
Recommender systems have become an integral part of e-commerce sites and other businesses like social networking, movie/music rendering sites. They have a huge impact on the revenue earned by these businesses and also benefit users by reducing the cognitive load of searching and sifting through an overload of data. Recommender systems personalize customer experience by understanding their usage of the system and recommending items they would find useful.
In order recommend a product to a user, we need to understand the user and their preferences. this can be seen from the user interaction with the website. In this case we have user reviews from where we can analyse the user behaviour and work on models to provide personalised Recommendations to the user.

Distribution of reviews Over The years For 'A. Dent "Aragorn"



Grouping on category which we got in previous step and getting the count of reviews as shown in below plot.



Recommendation test

- Based on product reviews, for B00INNP5VU average rating is 4.045267489711934.
- The first similar product is B008B1125W av-

erage rating is 4.154696132596685.

- The second similar product is B005JACJ50 average rating is 4.52972972972973.
- Based on product reviews, for B00IVPU786 average rating is 4.7975460122699385.
- The first similar product is B009SK57HY average rating is 4.445783132530121.
- The second similar product is B005TUQV0E average rating is 4.638655462184874.

Accuracy of the model : predicting the accuracy of the model by comparing the number of true positives. and true negatives.

Mean Square Error : Measures average squared error of our predictions. For each point, it calculates square difference between the predictions and the target and then average those values.

	precision	recall	f1-score	support
3	0.40	0.50	0.45	34
4	0.92	0.88	0.90	208
accuracy			0.83	242
macro avg	0.66	0.69	0.67	242
weighted avg	0.84	0.83	0.83	242

Test Accuracy : 0.8264462809917356
Mean Square Error : 0.17355371900826447

Conclusion

- Performed Exploratory Data Analysis on dataset to come up with insights such as rating, review trends over the years, range of ratings every year, helpfullness of the reviews, price range of products.
- Feature Extraction using TF-IDF vectorization for both the problems.
- **Sentiment Analysis :** Given a product and its reviews and ratings along with helpfullness score, predict the sentiment of review, whether the reveiw is 'positive' or 'negative'.
- Logistic Regression is comes out to be the best model to classify the sentiment of reviews.
- **Recommender System :** Given a product and its reviews and ratings along with helpfullness score, recommended the next best product based on the review summary and the overall rating.
- For recommendations, Analysing most reviewed user help us to understand user behaviour. Applied K-Nearest Neighbours with k=2 is used to get 2 nearest products based on review summary and overall review score, recommend the products to the user.

Future Work

- Customize the recommender for remaining product categories in Amazon Review dataset would be the next step. This recommender makes use of ratings, summary of review given by users. Making use of review text given by users would be interesting as a further step.
- Analysing whole reviews and Predicting the sentiment.
- Finding more patterns in the data to make more accurate recommendations.
- apply advanced techniques like BERT, wor2Vec to generate more presonalized recommendations.