



IJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 Issue: IV Month of publication: April 2023

DOI: <https://doi.org/10.22214/ijraset.2023.50223>

www.ijraset.com

Call:  08813907089

E-mail ID: ijraset@gmail.com

Sentiment Analysis Tool for Amazon Product Reviews

Dangeti Venkata Sai Krishna¹, Deepika Sanapathi², Devi Sishla Reddy³, Dokkara Jaswanth⁴

^{1, 2, 3, 4}Department of CSE GITAM Deemed University Visakhapatnam, India

Abstract: Sentiment analysis has become a vital component of modern data analysis, particularly for businesses that rely on customer input to improve their products and services. We employ Natural Language Processing (NLP) techniques to analyze sentiment in Amazon product reviews in this study. Our major goal is to categorize the assessments based on whether they are good, negative, or neutral. We'll use Amazon product review data, which includes a large number of reviews from various categories, such as books, electronics, and clothes.

Keywords: Web Scraping, Sentiment Analysis, Amazon, Product Reviews, Positive, Negative, Neural, Beautiful Soup, Natural Language Tool Kit (NLTK), Flask, VADER.

I. INTRODUCTION

The project is based on Amazon product reviews since Amazon is one of the largest online marketplaces, with millions of products and billions of customer reviews. As a result, analyzing the mood of Amazon customer reviews may give valuable insight into how people act, what they prefer, and how they feel about certain things. We apply natural language processing algorithms for sentiment analysis on Amazon product reviews in this research.

The power of sentiment analysis to extract key insights from massive amounts of textual data has lately piqued the curiosity of many in natural language processing. A customer review or social media post, for example, might be the topic of a sentiment analysis, which categorizes the text's sentiment, attitude, and opinion as positive, negative, or neutral. Businesses want this data since it may help them better grasp the viewpoints of their customers and modify the quality of their goods and services.

The project outputs can assist the consumer and business in understanding the product better, whether it is to acquire a product or improve the product's experience.

II. PROBLEM IDENTIFICATION

When discussing online shopping, we still determine what to buy and which to buy. It's always difficult when we can't pinpoint the specific reason for online purchases since we aren't sure what to buy in the first place. Even if we opt to buy anything, various elements come into play, such as the user experience, after-sale services, upgrades and alterations made after the sale, and so on. The average consumer takes two to three hours to decide what to buy. Furthermore, one may compare the items they wish to buy, the optimum time to buy them, and whether or not there are any sales.

These features are also stated in consumer reviews who have used the purchased item. Furthermore, developing an idea by reading each review takes time and work. We created the notion of creating a review analysis based on numerous reviews to generate a summary that gives an overview and an idea to make a quick choice to reduce the time necessary to make a purchase. These elements enticed consumers to write and read reviews on online shopping websites after they purchased the items and wanted to share their experiences. Because it has been demonstrated that reviews influence the majority of transactions, they are critical to our mission.

After making choices of features, we decided to take the product review page link as input; this helps us to speed up the process of web scraping, as we could potentially extract the reviews faster when we can give the link directly without compromising execution speeds. And the problem in giving the product page directly is that we cannot determine which page directly, and there are several pages for the reviews on the same product, and the link looks something like this : [\[https://www.amazon.in/Product Name/product-reviews/ASINValue/ref=id/ie=UTF8reviewerType=allreviews\]](https://www.amazon.in/Product Name/product-reviews/ASINValue/ref=id/ie=UTF8reviewerType=allreviews)

And this is not ideal for this use case. Instead, we suggest using the link format as :

[\[https://www.amazon.in/ProductName/product-reviews/ASIN Value/ref=id/ie=UTF8reviewerType=allreviewspageNumber=x\]](https://www.amazon.in/ProductName/product-reviews/ASIN Value/ref=id/ie=UTF8reviewerType=allreviewspageNumber=x)

This way, we can target the loop used for extracting multiple pages simultaneously, with the variable ['x'] as the number of pages we wish to extract using the program.

III. OBJECTIVES

The primary goal of this study is to gain a better understanding of emotion towards a product as represented through text reviews. People, it is assumed, will not spend much time comprehending and comparing the text reviews one by one.

A. Determining The Overall Sentiment

One of the key goals of sentiment analysis is to evaluate whether a product review's sentiment is favorable, negative, or neutral. This can assist firms in gauging consumer attitudes towards their product and identifying areas for development.

B. Why Do We Need To Identify Key Topics

Sentiment analysis may also assist in identifying the important subjects that customers discuss in their product evaluations. Businesses may better learn which product features are most important to customers and how they might enhance those qualities by analyzing the sentiment of reviews addressing certain subjects.

C. Analysis Of Sentiment Is An Essential Parameter To Monitor Customer Satisfaction

Businesses may monitor changes in consumer satisfaction and discover possible concerns early on by analyzing the sentiment of product evaluations over time. This can assist firms in changing their product or service before unfavorable feedback spreads.

D. The Comparison Could Also Be Against Competitors

Sentiment analysis may also compare a company's goods to its rivals. Businesses can find areas where they may have a competitive edge or disadvantage by analyzing the sentiment of product evaluations for numerous items.

E. How Do We Improve Marketing With These Results

Sentiment analysis may also reveal information about the words and phrases people use to describe a product. This can assist organizations in improving their marketing messaging and connecting with their target audience.

IV. IMPLEMENTATION

Data from Amazon product reviews were obtained via web scraping. For this case, we used the BeautifulSoup package to read and extract HTML pages. The proceeding step is to clean and prepare the data by removing superfluous information such as URLs, HTML components, and non-alphanumeric characters. Stop words should also be eliminated, and words should be stemmed/lemmatized to standardize the language. We used the NLTK package, which has stemmer and tokenization techniques that help us to stem and tokenize text. And this provides us with guidelines for categorizing the reviews as positive, negative, or neutral, we used the VADER package, which has a sentiment analyzer that helps us to generate scores for the individual sentence, as in this case, it is the review titles. The review titles retrieved for that product are the data utilized for sentiment analysis.

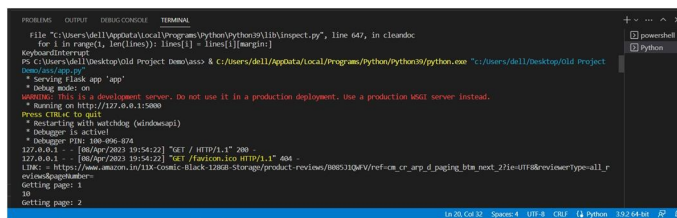


Fig. 1. Flask implementation in background

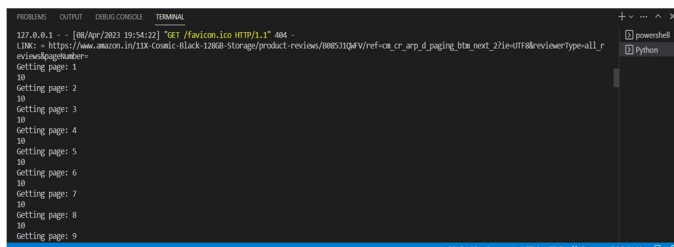
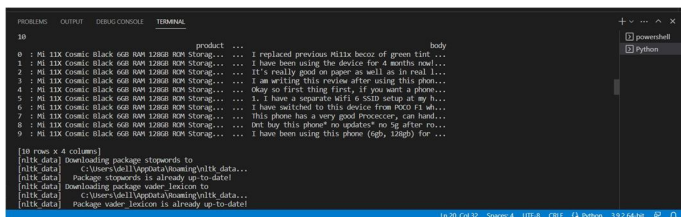


Fig. 2. The input link taken as to extract reviews

Visualize the sentiment analysis data using pie charts to gain insights into the overall sentiment of the product reviews. When depicted the pie chart depicts the total ratings for the retrieved pages per review, which are likewise calculated in percentage terms. This is achieved using pandas, plottypy, and Seaborn libraries to generate pie charts in a visually appealing format. Consequently, the outcomes area representation and analyzed version of those data, which assists us in identifying potential for product or marketing plan improvement. This sentiment analysis procedure may be automated for new Amazon product reviews, and the sentiment analysis tool can be deployed to a web application. As a result, any potential consumer can utilize it in the future.

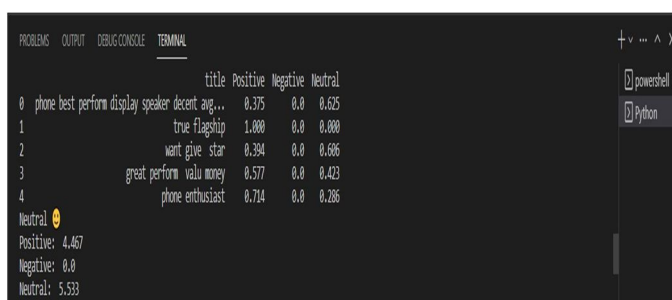


```

10 product ... body
0 : MI 11X Cosmic Black 6GB RAM 128GB ROM Storage... I replaced previous Mi11x becoz of green tint...
1 : MI 11X Cosmic Black 6GB RAM 128GB ROM Storage... I have been using the device for 4 months now...
2 : MI 11X Cosmic Black 6GB RAM 128GB ROM Storage... IT'S really good on paper as well as in real l...
3 : MI 11X Cosmic Black 6GB RAM 128GB ROM Storage... I am writing this review after using this phon...
4 : MI 11X Cosmic Black 6GB RAM 128GB ROM Storage... Okay so first, thing first, if you want a pho...
5 : MI 11X Cosmic Black 6GB RAM 128GB ROM Storage... 1. I have a separate MFI 6 SSD setup at my h...
6 : MI 11X Cosmic Black 6GB RAM 128GB ROM Storage... I have switched to this device from POCO P3 ab...
7 : MI 11X Cosmic Black 6GB RAM 128GB ROM Storage... this phone has a very good processor, cam hand...
8 : MI 11X Cosmic Black 6GB RAM 128GB ROM Storage... But hey this phone is update in 5g after rou...
9 : MI 11X Cosmic Black 6GB RAM 128GB ROM Storage... I have been using this phone (6gb, 128gb) for ...

[10 rows x 4 columns]
[nltk_data] downloading package stopwords to
[nltk_data] C:\Users\Yash\AppData\Local\nltk_data...
[nltk_data] Package stopwords is already up-to-date!
[nltk_data] downloading package vader_lexicon to
[nltk_data] C:\Users\Yash\AppData\Local\nltk_data...
[nltk_data] Package vader_lexicon is already up-to-date!
  
```

Fig. 3. The final overview of results extracted shown in terminal



```

title Positive Negative Neutral
0 phone best perform display speaker decent ag... 0.375 0.0 0.625
1 true flagship 1.000 0.0 0.000
2 want give star 0.394 0.0 0.606
3 great perform valu money 0.577 0.0 0.423
4 phone enthusiast 0.714 0.0 0.286

Neutral 😊
Positive: 4.467
Negative: 0.0
Neutral: 5.533
  
```

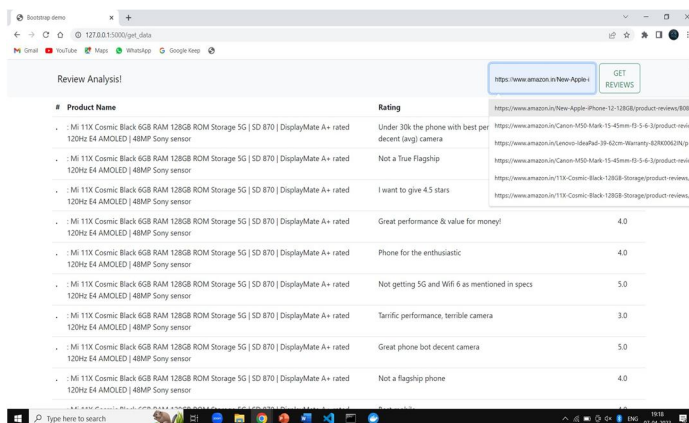
Fig. 4. The raw output scores for individual sentiment

It should be noted that these processes are not necessarily linear and that iteration may be required to increase the model's performance. Several tools and frameworks, such as NLTK, are also available to aid the implementation process.

V. RESULTS

The result that is generated using this tool are as follows, and these results include an overview of reviews extracted and following up with the results of the sentiment analysis for the review title and the comparison of ratings(out of 5) in the form of a pie chart. And this pie chart is in form with equivalent percentages of those ratings.

The review titles are the inputs considered by this senti- mental analysis tool because that is one parameter where the user tries to give short phrases for what he is trying to express. Generating individual sentiment scores for those titles can give us an overview of what the user is probably thinking. And after giving the title, he/she expresses thoughts and overviews his thoughts regarding the product itself. Hence we decided to take this title as input for the tool to generate sentimental analysis and display appropriate results for understanding.



#	Product Name	Rating
0	MI 11X Cosmic Black 6GB RAM 128GB ROM Storage 5G SD 870 DisplayMate A+ + rated 120Hz E4 AMOLED 48MP Sony sensor	Under 30k the phone with best per decent ag camera
1	MI 11X Cosmic Black 6GB RAM 128GB ROM Storage 5G SD 870 DisplayMate A+ + rated 120Hz E4 AMOLED 48MP Sony sensor	Not a True Flagship
2	MI 11X Cosmic Black 6GB RAM 128GB ROM Storage 5G SD 870 DisplayMate A+ + rated 120Hz E4 AMOLED 48MP Sony sensor	I want to give 4.5 stars
3	MI 11X Cosmic Black 6GB RAM 128GB ROM Storage 5G SD 870 DisplayMate A+ + rated 120Hz E4 AMOLED 48MP Sony sensor	Great performance & value for money!
4	MI 11X Cosmic Black 6GB RAM 128GB ROM Storage 5G SD 870 DisplayMate A+ + rated 120Hz E4 AMOLED 48MP Sony sensor	Phone for the enthusiastic
5	MI 11X Cosmic Black 6GB RAM 128GB ROM Storage 5G SD 870 DisplayMate A+ + rated 120Hz E4 AMOLED 48MP Sony sensor	Not getting 5G and Wifi 6 as mentioned in specs
6	MI 11X Cosmic Black 6GB RAM 128GB ROM Storage 5G SD 870 DisplayMate A+ + rated 120Hz E4 AMOLED 48MP Sony sensor	Tariffic performance, terrible camera
7	MI 11X Cosmic Black 6GB RAM 128GB ROM Storage 5G SD 870 DisplayMate A+ + rated 120Hz E4 AMOLED 48MP Sony sensor	Great phone bot decent camera
8	MI 11X Cosmic Black 6GB RAM 128GB ROM Storage 5G SD 870 DisplayMate A+ + rated 120Hz E4 AMOLED 48MP Sony sensor	Not a flagship phone

Fig. 5. Overview of Reviews titles and Ratings

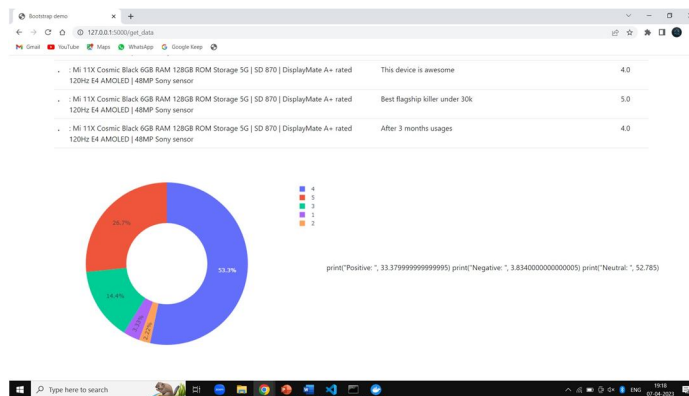


Fig. 6. Final Output representation

By portraying the rating's comparison individually represented as percentages also helps the user to conclude as this is a better representation of ratings than the usual representation of individual bars on the Amazon review page.

VI. CONCLUSIONS AND FUTURE SCOPE

Based on the analysis of Amazon product reviews, several conclusions can be drawn and future scope for review analysis: Customers are more likely to post longer and more descriptive evaluations for things they like or dislike. Negative reviews frequently emphasize product flaws, whereas favorable reviews highlight the product's qualities and benefits. Customers are more likely to trust and buy items with high average ratings and many favorable reviews. The language used in reviews might provide information about a customer's feelings, opinions, and preferences. Sentiment analysis may assist firms in identifying trends and patterns in consumer feedback, allowing them to make data-driven decisions.

Deep learning and machine learning are advanced natural language processing techniques that may be used to increase the accuracy of sentiment analysis and extract additional insights from reviews. Multilingual analysis may be used to comprehend client input from many languages and geographies. But it is not focused as of now in this project since that needs more sophisticated techniques to understand even for analytical purposes. To acquire a more thorough picture of consumer preferences and behavior, review analysis can be integrated with other data sources such as sales data and social media data. Visualization technologies may be used to develop interactive dashboards that provide review data to decision-makers in an easy-to-understand way.

REFERENCES

- [1] Fankar Armash Aslam, Hawa Nabeel Mohamed, "Efficient Way Of Web Development Using Python And Flask," International Journal of Advanced Computer Research, January 2015.
- [2] Vidhi Singrodia, Anirban Mitra, Subrata Paul, "A Review on WebScraping and its Applications," 2019.
- [3] Anil Singh Parihar, "A Study on Sentiment Analysis of Product Reviews," International Conference on Soft-computing and Network Security (ICSNS), 2018.
- [4] Dilesh Tanna, Manasi Dudhan, "Sentiment Analysis on Social Media for Emotion Classification," 19 June 2020.
- [5] Binju Saju, Siji Jose, "Comprehensive study on sentiment analysis: Types, approaches, recent applications, tools and APIs.," Federal Institute of Science and Technology, Mookanoor, Kerala, July 2020.
- [6] Zoubir AM, Iskander DR, "Bootstrap methods and applications [from the guest editors]," IEEE Signal Process Mag., 2007.
- [7] Venus Khim-Sen Liew, "An overview on various ways of bootstrap methods," University Malaysia, Sarawak, January 2008. .



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)