

# Product Rank Algorithm Along with Sentiment Analysis on Reviews of E-Commerce Websites

Suyash Lokhande<sup>1</sup>, Vedant Sharma<sup>2</sup>, Sarvesh Dhakunkar<sup>3</sup>, Suyash Chopda<sup>4</sup>, Dr. B.S. Satpute<sup>5</sup>,

<sup>1</sup>Department of Computer Engineering, Dr. D.Y. Patil Institute of Technology, Pimpri, Pune, Maharashtra, India

<sup>2</sup>Department of Computer Engineering, Dr. D.Y. Patil Institute of Technology, Pimpri, Pune, Maharashtra, India

<sup>3</sup>Department of Computer Engineering, Dr. D.Y. Patil Institute of Technology, Pimpri, Pune, Maharashtra, India

<sup>4</sup>Department of Computer Engineering, Dr. D.Y. Patil Institute of Technology, Pimpri, Pune, Maharashtra, India

<sup>5</sup>Department of Computer Engineering, Dr. D.Y. Patil Institute of Technology, Pimpri, Pune, Maharashtra, India

\*\*\*

**Abstract** - Everyone's day to day lives are tormented by COVID-19. People round the world are asked to remain at home to slow down the pandemic, that they'd never anticipated or experienced before. Shoppers have started realizing that this could be the 'New Normal' and have begun ever-changing their lifestyles consequently but are still afraid of stepping outside their homes even to buy for the essential product. Therefore, in several families, online grocery, apparel, and entertainment searching are being done using varied e-commerce websites. The fundamental purpose of our analysis is to supply consumers with the optimum product by scrutiny them on various product retail sites using a product ranking algorithm. This system analyses and provides the most affordable nonetheless very best quality products to the buyer. It not only ranks product however also considers reviews of consumers on every product and present a sentiment analysis approach on these reviews, that involves quantification of algorithmic program of thousands of reviews. And finally, providing the end-user summarized information regarding the expressed emotions within the means of intuitive and simple to understand visualization technique.

**Key Words:** Product Ranking Algorithm, Sentiment Analysis, Naïve Bayes, WordCloud, web scrapping, optimization

Therefore, it is necessary to filter supernumerary and expensive products and supply customers with few quality products offered on websites. Not solely numerous products however additionally heap of reviews are there for every product. Today, the common client depends heavily on user reviews and opinions expressed on product before making a choice to shop for. it's getting to be either positive or negative towards the merchandise. Therefore, reading all those reviews and so deciding to shop for a product is incredibly tedious and time overwhelming. This time will be saved by developing a system that automatically offers a best product of selection and save lot of your time. It's time to create machine to try to what Humans do currently. during this age of accelerating machine learning primarily based algorithms reading thousands of reviews to know a product is quite time consuming wherever we will polarize a review on specific category to know its quality among the patrons all over the globe. because the growing expenditure individuals are also involved concerning the costs as there's no-one from whom they'll negotiate for the costs. Due to these individuals continue adding product in their listing and waiting until the value is born is barely the choice. The projected system merges all the product offered at completely different websites and filters out the most effective product for the client. using internet crawlers, solely those products are offered to a client who comes underneath predefined parameters like worth, brand, product discount and popularity of the merchandise..

## 2. MODEL ARCHITECTURE AND TERMINOLOGY USED

### 2.1 System Architecture Diagram

The system is a Real-Time Search Engine, where users will find the best products among multiple e-commerce websites. The system searches the product over the website and is trained to perform web-scrapping continuously on the E-commerce website in order to extract all the information related to the product and the optimized list of the desired product will be displayed. It is necessary to trim millions of products on websites carefully. The parameters chosen for our web scrappers will decide the overall performance of the System. Few critical parameters on which a product will depend are the price of the product, brand and recognition of the product. Moreover, it will also help the user by giving a direct analysis of reviews using proper visualization

## 1. INTRODUCTION

E-commerce in India is the quickest growing platform where nearly over thousands of products are purchased in every second. Indian e-commerce is growing at an annual rate of 51, the best in the world, and is predicted to leap from \$3.535 trillion in 2019 to \$6.542 trillion in 2023. firms like Amazon, Flipkart, Myntra and Snapdeal makes trillions of profits once a year and most of them are during joyous seasons. At every web site, there are millions of products are offered and customers are not able to seek out product that are very offered at cheap costs, with additional discount and highest quality. Even filters fail to sort the product from low-high rate due to massive and dynamic set of information.

techniques. Users will be able to buy the product via the provided link of product from E-commerce websites.

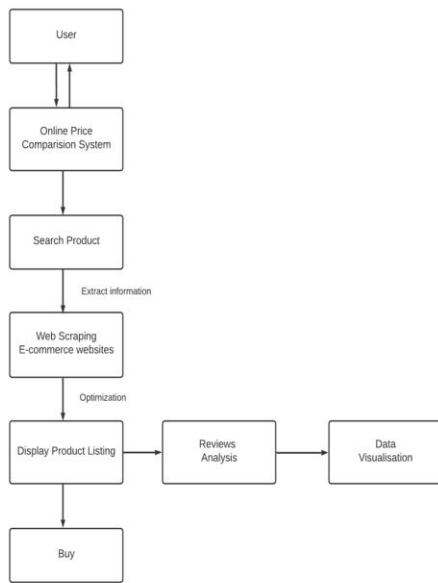


Figure 1- System Architecture

## 2.2 Algorithms or Terminology

### 1.Product Ranking Algorithm

This algorithm will create a value for each product which shows buying probability of product. For example:

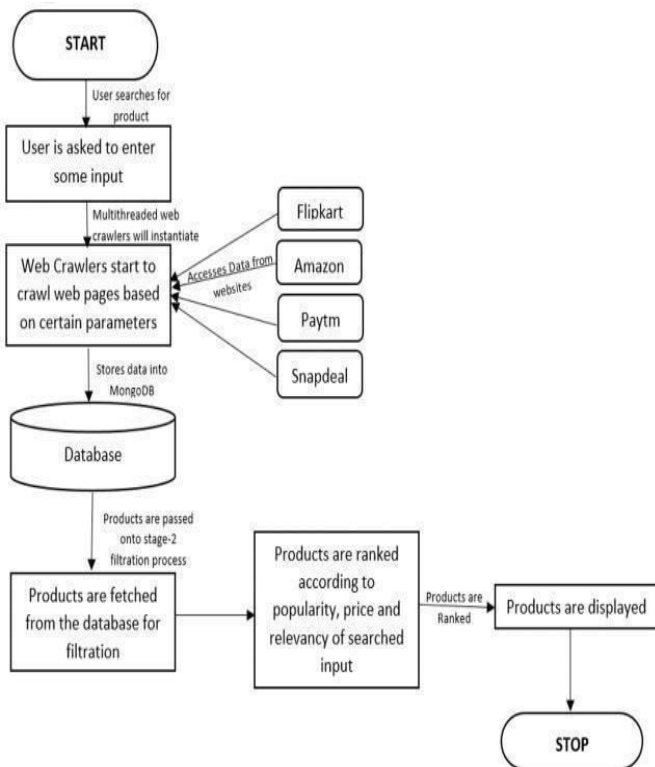


Fig 2. Product Ranking Algorithm

Let’s assume that our users will make their decision for purchasing a product only on the basis of its price and see if our machine learning model is able to learn such function. For this data set the price of movies will be ranging from 0 to

10, so we can artificially define the buy probability as follows:

$$\text{data}[\text{‘buy\_probability’}] = 1 - \text{data}[\text{‘price’}] * 0.1$$

In this system this algorithm is ranking product by considering all the parameters such as price, popularity, ratings and review count. Product rank algorithm will rank product with fewer prices and best quality product on top.

- Step 1: User enters input into the system that has to be searched.
- Step 2: Multithreaded web crawler activated and in parallel product searched on e-commerce websites taking user’s input as key parameter. This is Level-1 filtration process where the entire unnecessary and irrelevant product won’t be considered.
- Step 3: All the details of the selected products are stored by web crawlers into the CSV file and database.
- Step 4: All the scraped products passed through Product Rank algorithm (Level-2) where ranking to the product starts.
- Step 5: If the product comes under parameters predefined as then the rank of the product increases else decreases.
- Step 6: Items will be listed in decreasing order according to the product rank.
- Step 7: After collecting all the products, product along with its ranking will be displayed to the user.

	A	B	C	D	E	F	G	H
1	Product Name	Pricing(Rs.)	Ratings	Popularity	Reviews	Ram(GB)	ROM(GB)	Brand
2	Redmi Note 7 Pro (Space Black) 128 GB)	13999	4.5	276808	23205	6	128	Redmi
3	Mi A3 (More Than White) 128 GB)	14999	4.3	8692	766	6	128	Mi
4	Redmi Note 7 Pro (Space Black) 64 GB)	12999	4.5	276808	23205	6	64	Redmi
5	Redmi Note 6 Pro (Black) 64 GB)	11999	4.5	120346	10622	6	64	Redmi
6	Redmi K20 Pro (Carbon Black) 128 GB)	26999	4.5	55946	7011	6	128	Redmi
7	Redmi Note 8 (Neptune Blue) 128 GB)	14999	4.4	9497	759	6	128	Redmi
8	Redmi Note 8 (Moonlight White) 128 GB)	15995	4.4	9497	759	6	128	Redmi
9	Redmi Note 8 (Space Black) 128 GB)	15975	4.4	9497	759	6	128	Redmi
10	Redmi Note 7 Pro (Neptune Blue) 128 GB)	13999	4.5	276808	23205	6	128	Redmi
11	Mi A2 (Red) 128 GB)	13999	4.1	2351	208	6	128	Mi
12	Redmi Note 7 Pro (Nebula Red) 128 GB)	13999	4.5	276808	23205	6	128	Redmi
13	Redmi K20 Pro (Flame Red) 256 GB)	29999	4.5	12356	1738	8	256	Redmi
14	Redmi Note 7 Pro (Astro Moonlight White) 128 GB)	13999	4.5	276808	23205	6	128	Redmi
15	Redmi K20 Pro (Flame Red) 128 GB)	26999	4.5	55946	7011	6	128	Redmi
16	Mi A2 (Red) 64 GB)	10599	4.2	31840	2749	4	64	Mi
17	Redmi K20 Pro (Carbon Black) 256 GB)	29999	4.5	12356	1738	8	256	Redmi
18	Redmi K20 (Glacier Blue) 128 GB)	24999	4.5	45392	4919	6	128	Redmi
19	Redmi K20 (Flame Red) 128 GB)	24999	4.5	45392	4919	6	128	Redmi
20	Redmi K20 (Carbon Black) 64 GB)	21999	4.5	45392	4919	6	64	Redmi

Fig 3. Scrapped data (before ranking)

	A	B	C	D	E	F	G	H	I
1	Product Name	Pricing(Rs.)	Ratings	Popularity	Reviews	Ram(GB)	ROM(GB)	Brand	buy_prob
2	Redmi Note 7S (Sapphire Blue) 64 GB)	11999	4.4	39492	33080	4	64	Redmi	0.9953392
3	Redmi Note 7 Pro (Space Black) 64 GB)	12999	4.5	276808	23205	6	64	Redmi	0.9658958
4	Redmi Note 7 Pro (Neptune Blue) 128 GB)	13999	4.5	276808	23205	6	128	Redmi	0.9558958
5	Redmi Note 7 Pro (Nebula Red) 128 GB)	13999	4.5	276808	23205	6	128	Redmi	0.9558958
6	Redmi Note 7 Pro (Space Black) 128 GB)	13999	4.5	276808	23205	6	128	Redmi	0.9558958
7	Redmi Note 7 Pro (Astro Moonlight White) 128 GB)	13999	4.5	276808	23205	6	128	Redmi	0.9558958
8	Redmi Note 5 Pro (Gold) 64 GB)	13999	4.5	253476	25240	6	64	Redmi	0.9555976
9	Redmi Note 6 Pro (Black) 64 GB)	11999	4.5	120346	10622	6	64	Redmi	0.9476666
10	Mi A2 (Red) 64 GB)	10599	4.2	31840	2749	4	64	Mi	0.941943
11	Mi A2 (Red) 128 GB)	13999	4.1	2351	208	6	128	Mi	0.9014531
12	Redmi Note 8 (Neptune Blue) 128 GB)	14999	4.4	9497	759	6	128	Redmi	0.8957187
13	Mi A3 (More Than White) 128 GB)	14999	4.3	8692	766	6	128	Mi	0.8946452
14	Redmi Note 8 (Space Black) 128 GB)	15975	4.4	9497	759	6	128	Redmi	0.8859587
15	Redmi Note 8 (Moonlight White) 128 GB)	15995	4.4	9497	759	6	128	Redmi	0.8857587
16	Redmi K20 (Carbon Black) 64 GB)	21999	4.5	45392	4919	6	64	Redmi	0.8344682
17	Redmi K20 (Glacier Blue) 128 GB)	24999	4.5	45392	4919	6	128	Redmi	0.8044682
18	Redmi K20 (Flame Red) 128 GB)	24999	4.5	45392	4919	6	128	Redmi	0.8044682
19	Redmi K20 Pro (Flame Red) 128 GB)	26999	4.5	55946	7011	6	128	Redmi	0.7816156

Fig 4. Data After Ranking Product





will undergo Product Rank Algorithm where each product is going to be ranked accordingly. Filtration process during these stages is going to be highly rigorous, and this will make the best products available to the customers.

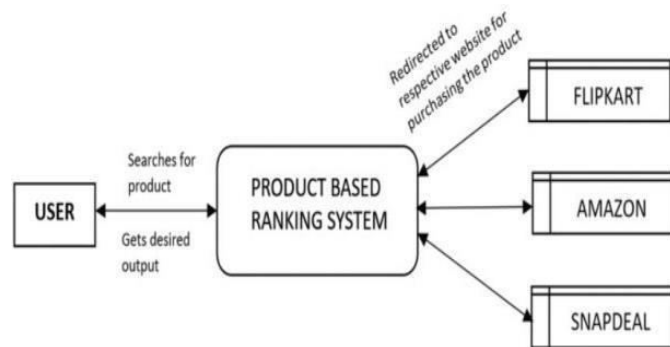


Fig 7. Product Based Ranking System

Figure shows the overview of the system. Users will search for products. The products will be scraped and processed in the system every hour to get updated results. When user decides the product for purchasing, control will then be redirected to the respective website for buying the product. Scraping of these products from the web site will occur one time every hour for an individual search entry. The following cases system will encounter:

*Case 1:*

Product is searched for the first time. Suppose n users are trying to search n different products at the same time then lots of computational power is needed for a practically impossible operation so before user searches for any product, the system will undergo maintenance every day and store the entries of those products in the database.

*Case 2:*

If the product is already in the database. If the product is already present in the database, price of the product will only be updated by web crawlers after every hour, which takes no time and will display it to the user. Since the scraped products are in thousands the amount of work done by web crawlers is comparatively less and can be achieved. Once scraping is done the necessary products will be filtered using Product Rank Algorithm. Product Rank Algorithm is going to perform ranking of products based on the product's popularity and lowest price of the product. It is working in the same way as that of Page Rank Algorithm in Google search engine to rank pages. Here the only difference, this system ranks each product based on how popular the product is on that day and what is the price of that product instead of web pages. This is the second stage filtration processes in which products having higher ranking displayed first.

As shown in Fig. 2 once user provides the desired input to system for searching, web crawlers start crawling websites to collect details of the desired product. Thereafter the result

obtained from search is stored into database in order to process it into product rank algorithm. After completion of computation, filtered products will be displayed to the user.

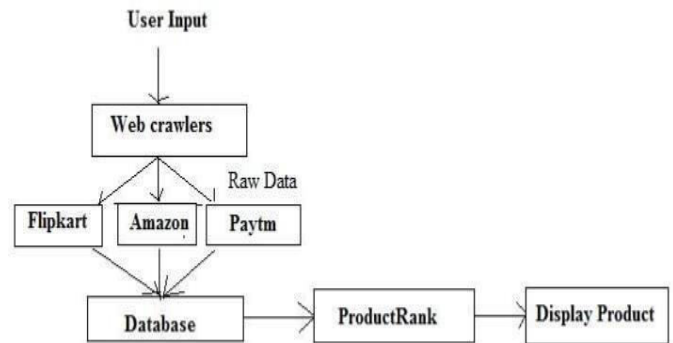


Figure. 8

**B. Sentiment analysis on reviews**

This paper focuses on mining reviews from the E-commerce websites like flipkart.com, which allows user to freely write the view. It automatically extracts the reviews of customers from the website. It also uses Word Cloud for performing sentiment analysis and for displaying the results. Following are the steps proposed for the data flow of the proposed system. The different components of the system for processing are as follows:

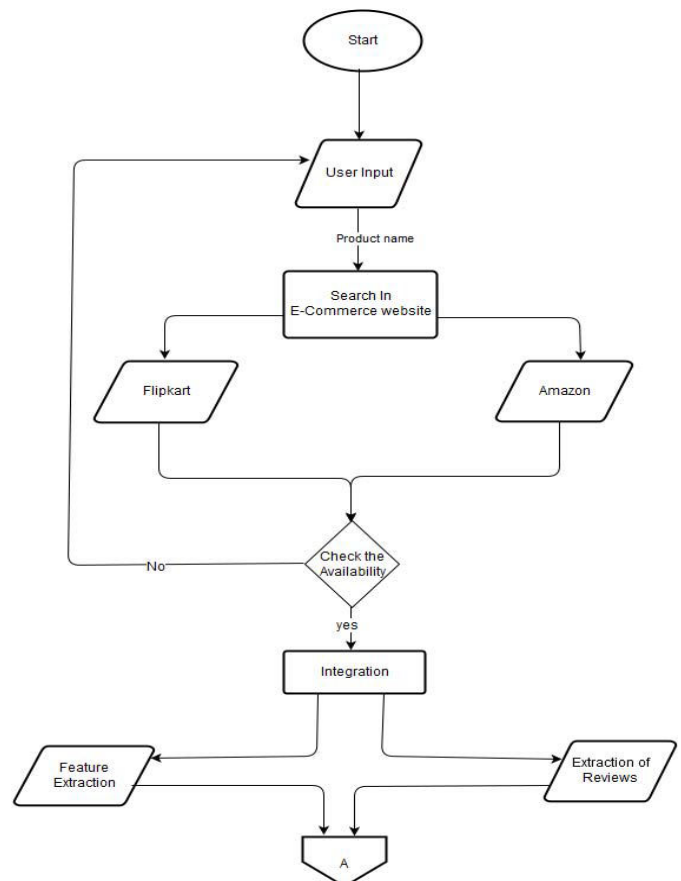


Fig-9 Sentiment Analysis

- User Input and Availability

The user enters the name of product that he desires. The name of the product entered by the user needs to be checked whether it is available in the E-Commerce site before proceeding with extraction.

- List of Product

This will display a list of products from which we have to pick a product of our choice to extract its review.

- Display Review List

This generates the dynamic link and displays all the reviews of the selected product.

- Analyzing Reviews

Here we perform analysis on the reviews that are being extracted using tools like Word Cloud and are then results are displayed on the screen.

#### 4. IMPLEMENTATION DETAILS

This work uses web scraping in Python and extracts the optimum products from a vast variety of products using a real time search engine. Products can also be verified by analyzing the reviews of those customers who have already bought the product. Python script is written mainly for these works.

##### A. System Description:

**Input:** The user will enter the name of the product in order to get the optimized list of products.

**Output:** Returns the optimized list of products along with analysis of the reviews of the product using proper visualization tool.

##### B. Software Requirements:

1. Operating System: Windows 10
2. IDE: PyCharm
3. Programming Language: Python and its supporting librarirs.
4. Chrome version: chrome version 64.0 or above

##### C. Hardware Requirements:

1. Hardware: intel core
2. Speed: 2.80 GHZ
3. RAM: 4GB
4. Key Board: Standard Windows Keyboard

##### D. User Interface

User of the system will be provided with Graphical User Interface to input the name of the product thet want to search or any text.

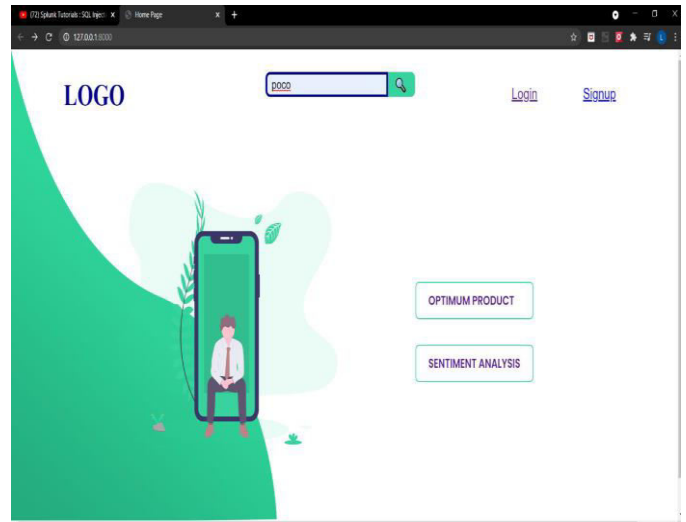


Fig.10: UI for input

After entering the name of the product, system returns the list of optimized product depending on various parameters. It also provides us the analysis of the reviews of that product by performing sentiment analysis on that product.

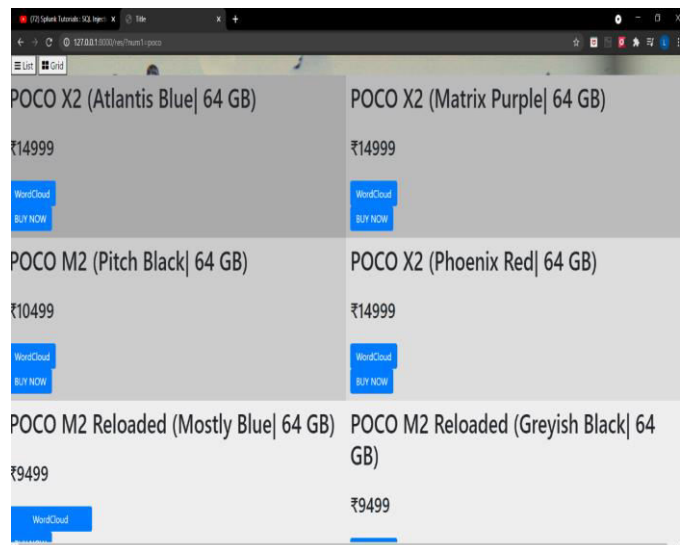


Fig.11 UI for optimized list

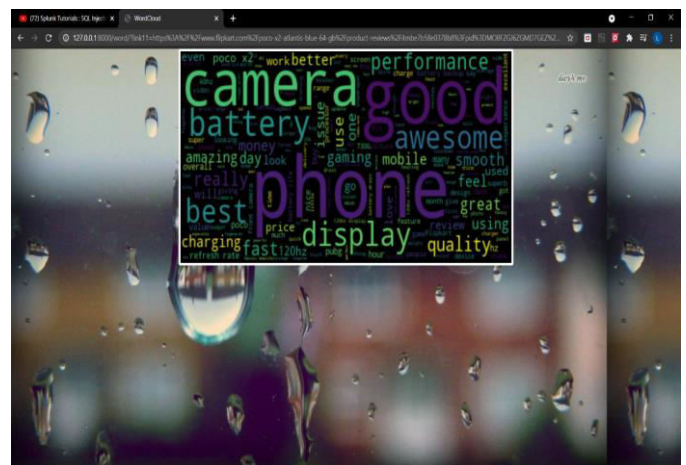
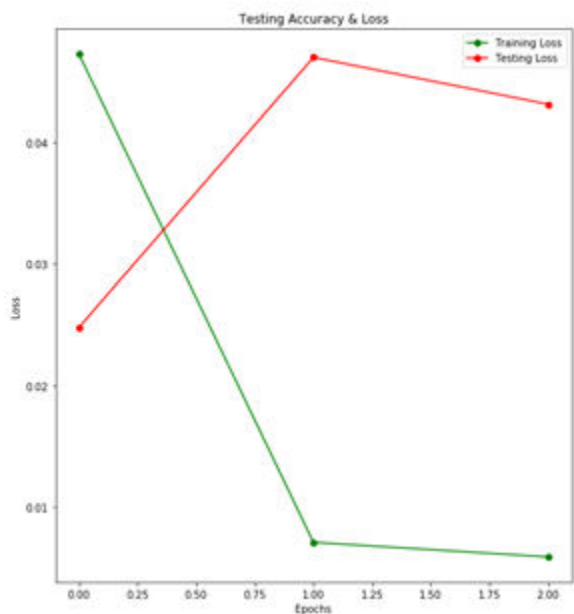
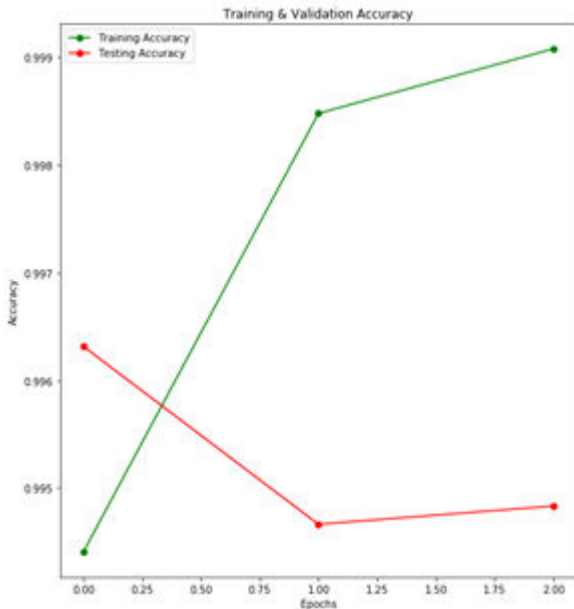


Fig.12 UI for Sentiment Analysis

## 5. RESULT AND DISCUSSION

### 5.1 Graphs

Below visualization showing the accuracy and loss values for both train and test sets over the epochs.



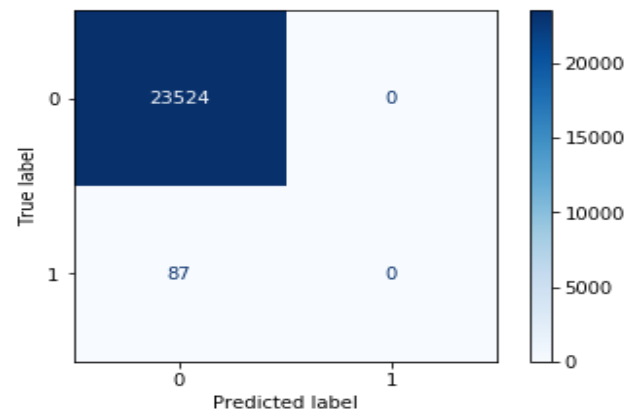
We see that the accuracy score remains roughly constant after 1 epoch, signifying that convergence has been achieved. We also note the similarity in results between train and test sets, signifying no overfitting.

Similar to the previous visualizing for accuracy, we note the loss stabilizes after the first epoch and is similar for both train and test sets.

### 5.2 Confusion Matrix

The confusion matrix for the test set is shown below. We can see overall four values

	precision	recall	f1score	support
0	1.00	1.00	1.00	23524
1	0.00	0.00	0.00	87
accuracy			1.00	23611
macro avg	0.50	0.50	0.50	23611
weighted avg	0.99	1.00	0.99	23611



## 6. CONCLUSION

Due to the increase of numerous products in e-commerce, customers are unable to find proper product or in many of the cases customers buy overpriced product which is non beneficial. In order to increase customers, experience at ecommerce, filtration of unnecessary products is needed and listings of quality products are needed and analyzing review to give one shot result is also needed. This system will provide customers the crystal-clear view of the searched product and provide cheapest product yet best quality product in one click.

## 7. FUTURE WORK

With the inrush of the big data era, the facts on the web have become way more dynamic, leading to better quality in Ranking and Sentiment Analysis. Therefore, as future work, we need to either design or use new product ranking algorithms that meet global standards of ranking challenges with efficiency and will provide users with both user specific as well as relevant information quickly and swiftly. Also, we can include local market sites along with retail sites like

Flipkart, Amazon, Myntra etc. to get better accuracy and usability. In this study, we used Word Cloud and Naïve Bayes for sentiment analysis without further modification. Still, in future work, it will be necessary to adjust tools and approaches to the sentiment analysis, particularly for mixed emotions, since our manual observations indicate that users actively discuss the pros and cons of a product. So, adding additional features, we can get the exact perception of the product by the user. And lastly, we will try to continue this research until we generalize it to all kinds of text-based reviews and comments.

[10]. Ashish Jain, Rajeev Sharma, Gireesh Dixit, Varsha Tomar, *Page Ranking Algorithms in Web Mining, Limitations of Existing methods and a New Method for Indexing Web Pages, International Conference on Communication Systems and Network Technologies, 978-0-7695-4958-3/13 IEEE DOI 10.1109/CSNT.2013.137 pp.640-645, 2013.*

#### REFERENCES

- [1]. <https://www.ibef.org/industry/ecommerce.aspx> 07th December 2017
- [2]. Sergey Brin and Lawrence Page, *The Anatomy of a Large Scale Hypertextual Web Search Engine. Computer Science Department, Stanford University, Stanford, CA 94305, USA, pp.1-20 1999.*
- [3]. Urvesh Rathod, Prof. Aruna Pavate, Vaibhav Patil. *Product Rank Based Search Engine for E-Commerce. 3rd International Conference for Convergence in Technology (I2CT), IEEE, Apr 2018 DOI: 10.1109/I2CT.2018.8529503*
- [4]. Prashast Kumar Singh, Arjit Sachdeva, Dhruv Mahajan, Nishtha Pande, Amit Sharma, *An approach towards feature specific opinion mining and sentimental analysis across e-commerce websites, 5th International Conference-Confluence The Next Generation Information Technology Summit (Confluenc),2014*
- [5]. <https://analyticsindiamag.com/step-by-step-guide-to-reviews-classification-using-svc-naive-bayes-random-forest/>
- [6]. Santhosh Kumar K L, Jayanti Desai, Jharna Majumdar, *Opinion Mining and Sentiment Analysis on Online Customer Review, IEEE International Conference on Computational Intelligence and Computing Research, 2016*
- [7]. Florian Heimerl, Steffen Lohmann, Simon Lange, Thomas Ertl, *Word Cloud Explorer: Text Analytics based on Word Clouds, 2014 47th Hawaii International Conference on System Science,2014*
- [8]. Vamsee Krishna Kiran M, Vinodhini R E, Archanaa R, Vimalkumar K, *User specific product recommendation and rating system by performing sentiment analysis on product reviews, 2017 International Conference on Advanced Computing and Communication Systems (ICACCS -2017), Jan. 2017*
- [9]. G.Angulakshmi, Dr.R.ManickaChezian, *An Analysis on Opinion Mining: Techniques and Tools. International Journal of Advanced Research in Computer and Communication Engineering, Vol. 3, Issue 7, July 2014.*