Sentiment Classifier on E-Commerce Product Reviews Using Logistic Regression

Prof. Pradeep N. Fale, Nikita Swain, Pranay Rahangdale, Nikhil Dhakate, Harshwardhan Bagde, Pranay Waghmare

[1]Assistant Professor, Department of Information Technology, Priyadarshini College of Engineering, Nagpur, India

[2]UG Scholar, Department of Information Technology, Priyadarshini College of Engineering, Nagpur, India

ABSTRACT

All the e-commerce sectors are dependent on product selling and engendering revenue via advertisements, promotions, and offers. The companies and brands fixate on analyzing reviews on the product by different buyers of different ages and countries. Through the machine learning techniques, it is now possible to analyze product review data of users of Amazon, Flipkart, Myntra. This paper is presented to show the conceptions and implementation to find sundry sentiments of customers buying products online. And with the avail of datasets, we will define the precision of the analysis.

Keywords - Sentiment Analysis, python, IOT, positive, negative, neutral

I. INTRODUCTION

In today's world the internet is the fastest growing thing and its horizon is becoming wider. In the shelter of internet connectivity, many online businesses have reached the level of success in just a few years. Online shopping is a trend these days, especially after the covid-19 pandemic. Online shopping platforms such as Amazon, Flipkart, Meesho, and Mantra are spreading their boom in the market because customers are buying products as well as sharing their reviews that help other customers to buy the same product from the same site. A brand becomes trending when people share their values and opinions.

In this paper, a machine learning-predicated sentiment analysis system is proposed to analyze the sentiments behind product reviews given by the customers of different products to relegate the desiderata with the utilization of natural language processing to extract, identify and characterize the sentiments on the substratum of reviews. These sentiment analysis are referred to as emotion mining and polarity classification and have a wide range of applications.

The E-commerce website mainly concentrate on positive feedback only. For sentiment analysis use to calculate the feedback comments in text. In which first it will calculate the positive and negative rating and then calculate overall rating and feature ratings and weights are figured based on accumulation of each. Sentiment analysis can be used in various form,

like can analyzed review of stock holders about particular company by using social media. To improve the business, businessman has to analyze the customer's reviews, and analyzing feedback data is necessary that will help in to find the customers actual requirement and existing product satisfaction. Sentiment analysis not only gives positive, negative and neutral but also analysis of happy, sad and not satisfied reviews. This fined the easiest way to collect the customer's opinion on products. Sentiment analysis can used to analyze the large amount of data formless data.

II. LITERATURE SURVEY

A DLMNN methodology is proposed aimed at SA of online products review and an IANFIS methodology is proposed aimed at future prediction of online products. The performance of both the proposed methodologies is analyzed. The proposed DLMNN is employed for three scenarios (GB, CB, and CLB) of RA. The comparison of those three scenarios for disparate numbers of data (from 1000 to 5000) concerning the performance measures of ps, rk, fs, and ac, is done. While comparing the '3' scenarios, the CLB scenario attains the best outcomes for product RA. And, while contrasting the IANFIS for future prediction against the existing ANFIS, the proposed IANFIS attains the highest values for ps, rk, fs, and ac.

P. Sasikala and L. Mary Immaculate Sheela et. al. [1]A DLMNN methodology is proposed aimed at SA of online products review and an IANFIS methodology is proposed aimed at future prediction of online product.
The performance of both the proposed methodologies is analyzed. The proposed DLMNN is employed for three scenarios (GB, CB, and CLB) of RA. The comparison of those three scenarios for disparate numbers of data (from 1000 to 5000) concerning the performance measures of ps, rk, fs, and ac, is done. While comparing the '3' scenarios, the CLB scenario

attain the best outcomes for product RA. And, while contrasting the IANFIS for future prediction against the existing ANFIS, the proposed IANFIS attains the highest values for ps, rk, fs, and ac. Hence, from the performance analysis, the paper infers that the proposed CLB scenario and IANFIS performed-well for SA and future prediction of online products. The system has a shortcoming such that the keyword processing only identifies the sentiment reflected in a particular word; it typically fails at providing all of the elements necessary to understand the complete context of the entire piece. In the future, the proposed system has been extended by solving the keyword processing problem and improve the performance using a hybridization algorithm in the future prediction process.

Shanshan Yi, Xiaofang Liu. Et, al., [2] This paper is motivated towards applying Machine Learning algorithms for learning, analysing and classifying the product information and the shop information based on the customer experience. The product data with customer reviews is collected from benchmark Unified computing system (UCS) which is a server for databased computer product lined up for evaluating hardware, support to visualization, software management. From the results and comparison, it has been found that machine learning algorithms outperform than other approaches. The proposed HRS system has higher values of MAPE which is 96% and accuracy is nearly 98% when compared to other existing techniques. Mean absolute error of proposed HRS system is nearly 0.6 which states that the performance of the system is significantly effective.

Anvar Shathik, Krishna Prasad K. et.al., [3] This paper summarizes the techniques for machine learning used in the analysis of emotions in the latest periods. Different application areas of sentiment analyze are also explored in such as industry, politics, public behavior, and finance. Through this paper, the impact of applying data transformations may improvise the

achievement of the methods of classification but the type of transformation lieson the dataset and the language it includes. Therefore, look at the details, make a selection of the features, apply transformations and filter the less relevant data making machine learning methods generalize and effective since the computers these days have limits and can't handle them all the data without prior review of any kind. The machine learning methods appear to typically offer the findings are identical and, again, depending on the form of those results. This paper assumes that applications of sentiment analysis will continue to grow in the future and that the implementation of sentiment analytical techniques will be standardized in various systems and services. The proposed future work will focus on three different characteristics chosen to investigate various datasets combining logistic regression and SVM algorithms. It can find unfair positive reviews and unfair negative reviews, reputation issues, and collusion and control through this work. The experimental method can study the accuracy, precision, and recall of both algorithms and can determine accurate and less time feature selection.

Alexander Ligthart et.al.,[4] The outcome of this tertiary study provides a comprehensive overview of the key topics and the different approaches for a variety of tasks in sentiment analysis. Different features, algorithms, and datasets used in sentiment analysis models are mapped. Challenges and open problems are identified that can help to identify points that require research efforts in sentiment analysis. In addition to the tertiary study, they also identified recent 112 deep learning-based sentiment analysis papers and categorized them based on the applied deep learning algorithms. According to this analysis, LSTM and CNN algorithms are the most used deep learning algorithms for sentiment analysis.

Jaspreet Singh et.al.,[5] Words and phrases bespeak the perspectives of people about products, services,

governments and events on social media. Extricating positive or negative polarities from social media text denominates task of sentiment analysis in the field of natural language processing. The exponential growth of demands for business organizations and governments, impel researchers to accomplish their research in sentiment analysis. This paper leverages four state-of-the-art machine learning classifiers viz. Naïve Bayes, J48, BF Tree and One R for optimization of sentiment analysis. The experiments are performed using three manually compiled datasets; two of them are captured from Amazon and one dataset is assembled from IMDB movie reviews.

Nirag T. Bhatt1, Asst. Prof. Saket J. Swarndeep2 et. al.,[6] In this article there are different machine learning techniques which are used for sentiment analysis. Mostly sentiment analysis done by using machine learning classifier like SVM (support vector machine), Random forest, Naïve Bayes. In this we are seeing some paper which are help new researcher to found a proper path for their new research. In this there is a proposed method of new research program. Social media is biggest medium to share people's opinion on different topics. Sentiment analysis uses machine learning technique and without any human interruption machine will give and accurate sentiment of the people. Sentiment analysis turn text into positive, negative or neutral. So, any company or foundation or movie reviewer can take the opinion of the people and take further steps according that.

Mirsa Karim,Smija Das. et.al.,[7] These days there is an expansion in review websites. It has turned out to be considerably more intricate to mine fundamental data from survey sites and take proper choice .Using Natural Language Processing, there is need to identify sentiment of content or document. In this paper Sentiment Analysis is done in view of Rule based mechanism and machine learning approach. Both of these strategies are analysed and discovered that machine learning is most appropriate for Sentiment

Analysis in light of the exactness measurement. Sentiment Vader and Sentiwordnet are the Rule based algorithms utilized and LDA analysis on Naive bayes is the machine learning strategy used.

[8] In this study, a sentiment analysis application for twitter analysis was conducted on 2019 Republic of Indonesia presidential candidates, using the python programming language. There are several steps taken to conduct this sentiment analysis, which is to collect data using libraries in python, text processing, testing training data, and text classification using the Naïve Bayes method. The Naïve Bayes method is used to help classify classes or the level of sentiments of society. The results of this study found that the value of the positive sentiment polarity of the Jokowi-Ma'ruf Amin pair was 45.45% and a negative value of 54.55%, while the Prabowo-Sandiaga pair received a positive sentiment score of 44.32% and negative 55.68%. Then the combined data was tested from the training data used for each presidential candidate and get an accuracy of 80.90% ≈ 80.1%. In this study a comparison was carried out using the naïve bayes, svm and K-Nearest Neighbor (K-NN) methods which were tested using RapidMiner by producing a naïve bayes accuracy value of 75.58%, svm accuracy value of 63.99% and K-NN accuracy value of 73.34%.

III. AIM & OBJECTIVES

- We aim to perform Sentiment Analysis of product based reviews.
- Scrapping product reviews on various websites featuring various products specifically amazon.com.
- Analyse and categorize review data. Analyse sentiment on dataset from document level (review level).
- Categorization or classification of opinion sentiment into - Positive, Negative, Neutral, etc.

- Data used in this project are online product reviews collected from "amazon.com", "flipkart.in", "myntra.com"

IV. METHODOLOGY

In this chapter all the methodologies are discussed along with diagrams and figures to show the actual working of the proposed system. The methodology consists of basic requirements, planning, process and implementation. Below given is a flow diagram of the text processing method.
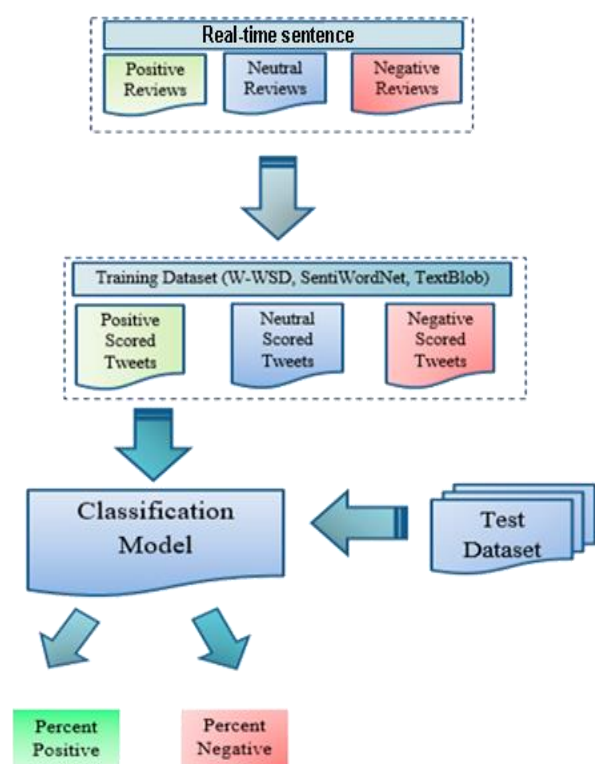

Fig. 1 Text Processing

There are several processes that are carried out in this text processing: firstly we collect data, in this study we using data tweets are collected from Twitter social media by using a crawler. Furthermore, we parse the sentence by describing it verbatim. Here in after, we do the tokenization process that is cleaning the tweet and selecting the meaningful words. Then, we do text mining using naïve bayes method, the process of text processing can be seen in Figure above:
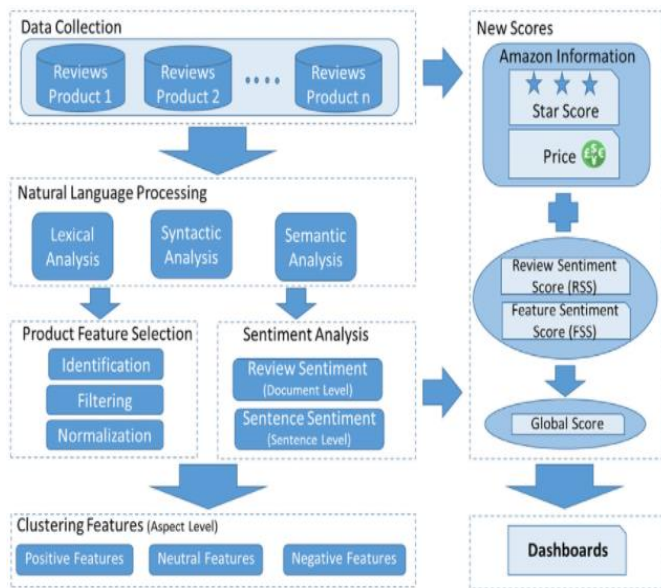
Figure. 2 Sentiment Analysis Process Flow using NLP

As shown above, in fig. 2, the process of collecting data from various product reviews of different e-commerce websites is mentioned. In this figure, we can see the complete analysis process with Natural Language Processing (NLP), which is also a base of this project. The complete process is described below:

Process 1: Firstly the data is collected of different products from various websites to start reviewing it.

Process 2: As the data, we have collected is in human language and needs to be converted first in binary codes. So the conversion process is done using NLP.

Process 3: Natural Language Processing shortly known as NLP will help to make natural human language usable to machine easel using NLTK known as Natural Language Toolkit.

Process 4: After the processing system will analyze what kind of ratings and reviews are affecting product selling with sentiments of the words.

Process 5: Analyzing the sentiments of the reviews will help us to understand the product quality and the satisfaction level.

V. PROPOSED METHODOLOGY

In this proposed system, Social media contain a large amount of raw data that has been uploaded by users. The data can be converted into valuable information by using sentiment analysis. Sentiment analysis is an approach that uses Natural Language Processing (NLP) to extract, convert and interpret opinion from a text and classify them into positive, negative or natural sentiment. Most of the previous study applied sentiment analysis into a product or movie review to better understand their customer and make the necessary decision to improve their product or services.

The feedback comments are considered as a source where buyers express their opinions more honestly and openly. The analysis of feedback comments on eBay and Amazon reveals that even if a buyer gives a positive rating for a transaction, she/he still leaves comments of mixed opinions regarding different aspects of transactions in feedback comments. Aspect opinion expressions, and their associated ratings (positive or negative) are first extracted from feedback comments. Dimension trust scores together with their weights are further computed by aggregating dimension ratings.

3.3 Algorithms & Tools

a) Random Forest

It is a supervised learning algorithm used in classification and regression problems. It is used to find result in accuray. It is also an effective algorithm when it comes to find classification of the images or text but this algorithm is specially designed for work on image producing not for text processing. But as a result, analysers and comparisons have a contribution in this project.

b) SVM Classifier

It is an algorithm generally used for clarification and it has its speciality in classification to find N number of features. SVM works very well and provides accuracy but the drawback of this algorithm is that it only works on limited datasets as we studied in refereed research papers. So this algorithm is used for result and accuracy comparison purposes only in this project.

c) Tkinter

It is an internal python library used to design GUI (Graphical User Interface) for softwares and projects front end. The GUI of this project is made using this tool, it is easy to implement and provides good looks to the system we made.

d) Kaggle

It is a dataset that provides data for testing and training modules, and helps users to build models in software. We have made use of this dataset to have more are more data for finding accuracy in the project.

e) E-commerce site dataset

Collection of product reviews datasets from various e-commerce websites for classification of results.

f) Machine Learning

It is a type of AI (Artificial Intelligence) which lets the application perform accuracy in obtaining results. With the help of these algorithms can be applied easily.

Comparison of Algorithms

SR.NO	ALGORITHM	ACCURACY
1.	NAÏVE BAYES	81.60
2.	SUPPORT VECTOR MACHINE	81.00
3.	K-NEAREST NEIGHBOUR	62.85
4.	LOGISTIC REGRESSION	84.20

In this paper,we have mainly used logical Regression Algorithm along with Naive Bayes Algorithm,Support vector Machine(SVM) , K-Nearest Neighbour(KNN), and Logistic Regression Algorithm. Generally, In previous related work they used atleast one or two algorithms at a time. But we proposed a work to do a combinely work along with four algorithms. By using all these four algorithms we proposed a comparative result to show by using different types of graphs, charts and tables.

Firstly , we had create different datasets of different types of products. Then we process it and show the comparative result. In our observation, we have found and recorded the accuracies of different algorithms. According to our observations, Naïve Bayes algorithm has an accuracy of 82.60%, Support Vector Machine (SVM) algorithm has an accuracy of 81.00%, K-Nearest Neighbour (KNN) algorithm provides the accuracy of 62.85%, And the Algorithm which is our prioritize work i.e. Logistic Regression Algorithm gives us the maximum accuracy of 84.20%. Also we had observed that when we take large number of datasets the accuracy difference of all four algorithms are large as compare to less number of datasets .

a) Naïve Bayes:-

Naïve Bayes algorithm used for text mining in process and it also assumes that features are conditionally independent. Actual datasets never be perfectly independent but they can be nearby. It has a higher bias but lower variance when it compared to logistic regression. When the datasets follows the bias then always the Naïve Bayes will be a superior classifier than other algorithm like logistic regression. It is easiest and fast so than it could make a prediction in real time.

b) Logistic Regression:-

Logistic regression is a classify algorithm which is used to understand dependent variable and independent variable, which is based on previous observation. It make an expected probability by a direct functional form while Naïve Bayes figures out how the data will generated on given the results. Logistic regression use for data analysis, when there are more classes, multi class.

c) K-Nearest Neighbour:-

K-Nearest Neighbour (KNN) is the vital algorithm in machine learning. KNN is the modest form of machine learning which can be classify the text. In KNN representation of k is the number of nearest neighbours will be used to classify the new data points. It is useful in prediction of wide variety of data.

d) SVM :-

Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms, which is used for Classification as well as Regression problems. However, primarily, it is used for Classification problems in Machine Learning.The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future.

Sentiment Analysis Algorithm:

Logistic regression predicts the probability of an outcome that can only have two values (i.e. a dichotomy). The prediction is based on the use of one or several predictors (numerical and categorical). A linear regression is not appropriate for predicting the value of a binary variable for two reasons: 1) a linear regression will predict

values outside the acceptable range (e.g. predicting probabilities outside the range 0 to 1). 2) Since the dichotomous experiments can only have one of two possible values for each experiment, the residuals will not be normally distributed about the predicted line.

On the other hand, a logistic regression produces a logistic curve, which is limited to values between 0 and 1. Logistic regression is similar to a linear regression, but the curve is constructed using the natural logarithm of the "odds" of the target variable, rather than the probability. Moreover, the predictors do not have to be normally distributed or have equal variance in each group. Logistic regression uses maximum likelihood estimation (MLE) to obtain the model coefficients that relate predictors to the target. After this initial function is estimated, the process is repeated until LL (Log Likelihood) does not change significantly.

$$\beta^1 = \beta^0 + [X^T W X]^{-1}.X^T(y-\mu)$$

- β is a vector of the logistic regression coefficients.
- W is a square matrix of order N with elements $n_i \pi_i (1 - \pi_i)$ on the diagonal and zeros everywhere else.
- μ is a vector pf length with elements $\mu_i = n_i \pi_i$
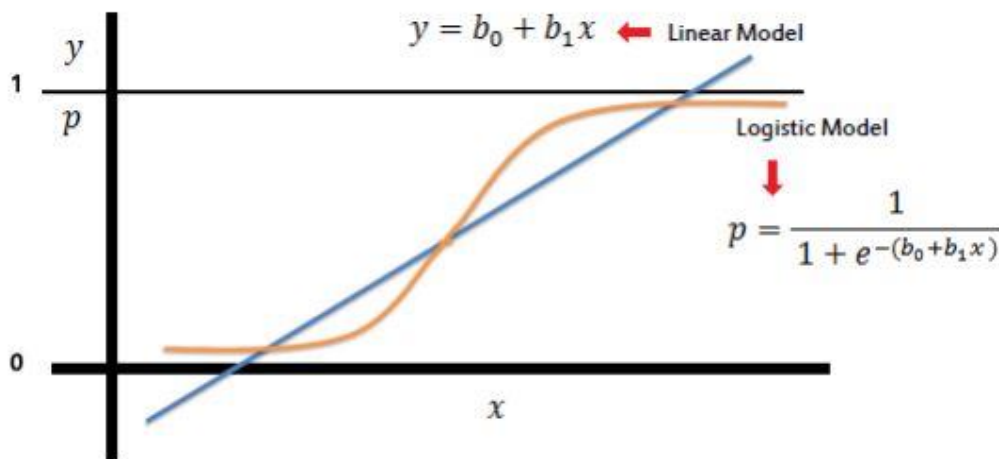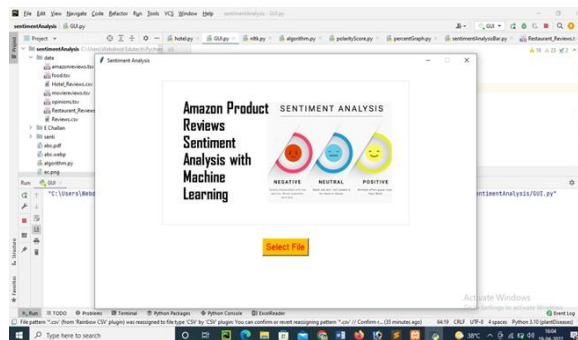

VI. RESULT AND DISCUSSION

fig 1. GUI

System GUI, With the vast amount of consumer reviews, this creates an opportunity to see how the market reacts to a specific product .
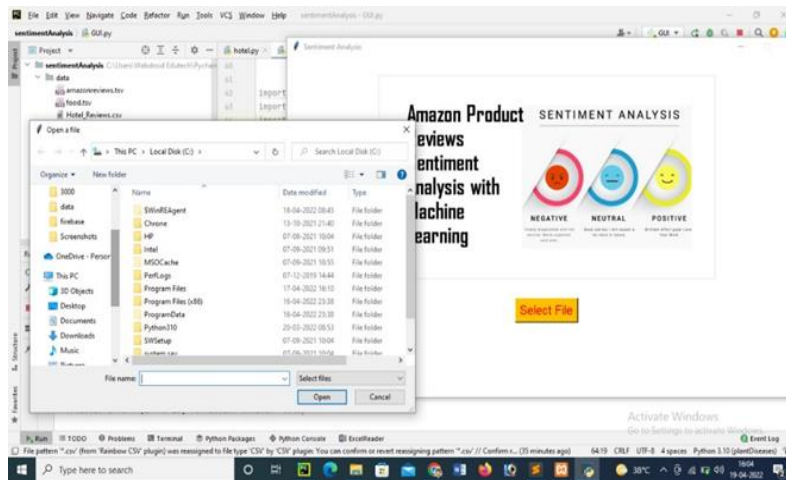

fig 2. Browse the File

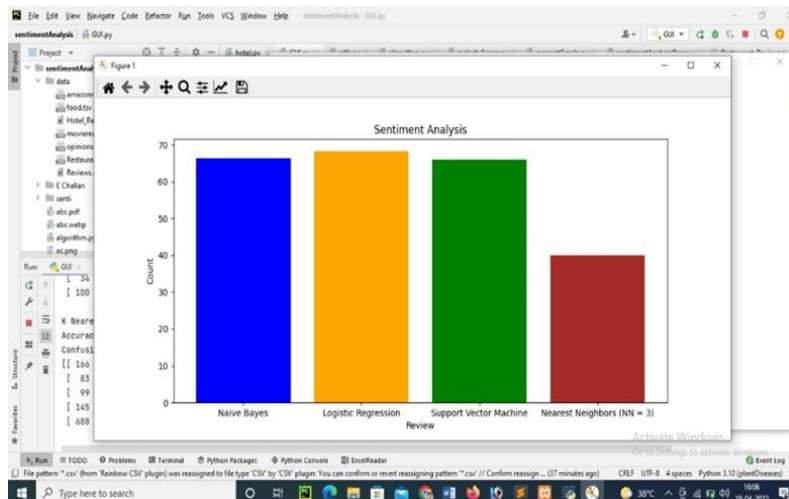Click on Select File Button, it will browse the file from system

Fig 3 Data Analysis Using

(Naive Bayes, Logistic Regression, SVM, Nearest Neighbor)Algorithm

Review of Data analysis using Naive Bayes, Logistic Regression, SVM, Nearest Neighbor Algorithm. It will analyze the ratings of data..
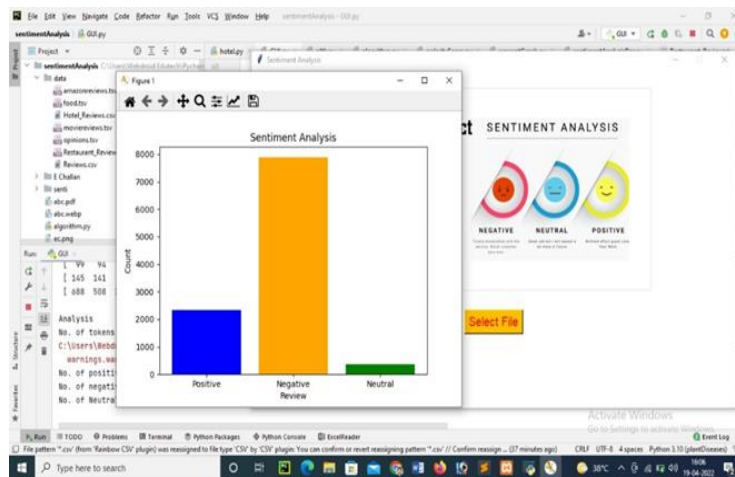
fig 4. Data classifier Review

Using the features in place, the data classifier that can determine a review's sentiment.

VII.CONCLUSION

Analyzing sentiments behind product review is as important as selling a product online without promotions as the review itself does the promotion for various products. Finding sentiments of the review helps to improve productivity and product selling in the market. This system is able to perform analysis of various e-commerce product reviews and shows the accuracy using random forest algorithm which is also a supervised machine learning technique. This will help to take steps toward bad product selling and circulate the actual reason behind reviews. It will used to evaluate the opinions of stock holder on particular product that is directly impact on company that will improve the performance of company, according to requirement of customer need that is customer satisfaction product. When there will be large amount of data to analyze the sentiment analysis help to give quick calculation. Product sentiment in customer feedback.

VIII. REFERENCES

[1]. P. Sasikala and L. Mary Immaculate Sheela Sentiment analysis of online product reviews using DLMNN and future prediction of online product using IANFIS (June 2020)

[2]. Shanshan Yi, Xiaofang Liu, Machine learning based customer sentiment analysis for recommending shoppers, shops based on customers' review (June 2020), Published online springer.

[3]. Anvar Shathik , Krishna Prasad K. A Literature Review on Application of Sentiment Analysis Using Machine Learning Technique(June 2020) International journal of Applied Engineering and Management Letters

[4]. Alexander Ligthart, Cagatay Catal , Systematic reviews in sentiment analysis: a tertiary study (Feb 2021) on customers' review (June 2020)

[5]. Jaspreet Singh, Gurvinder Singh and Rajinder Singh,Optimization of sentiment analysis using machine learning classifier (May 2017) Singh et al. Hum. Cent. Comput. Inf. Sci. (2017)

[6]. Nirag T. Bhatt1, Asst. Prof. Saket J. Swarndeep2 Sentiment Analysis using Machine Learning Technique (Dec 2020) International Research Journal of Engineering and Technology (IRJET)

[7]. Mirsa Karim,Smija Das."SENTIMENT ANALYSIS ON TEXTUAL REVIEWS "(March 2018) IOP Conf. Series: Materials Science and Engineering

[8]. F. A. Pozzi, E. Fersini, E. Messina and B. Liu, in Sentiment Analysis In Social Network, United States, Todd Green, 2017, p. 228.

[9]. S. Widoatmodjo, in Cara Cepat Memulai Investasi Saham Panduan Bagi Pemula, Jakarta, Kompas Media, 2012, p. 139.

[10]. Rajput, D. Singh, Thakur, R. Singh, Basha and S. Muzamil, Sentiment Analysis and Knowledge Discovery in Contemporary Business, United States of America: IGI Global, 2018.

[11]. C. A. Haryani, H. Tihari, Marhamah and Y. A. Nurrahman, Sentimen Analisis Kepuasan Pelanggan E-commerce Menggunakan Lexicon Classification dengan R, in Konferensi Nasional Sistem Informasi, Pangkalpinang, 2018.

[12]. A. Deviyanto and M. D. Wahyudi, Penerapan Analisis Sentimen Pada Pengguna Twitter Menggunakan Metode KNearest Neighbor, JISKa (Jurnal Informatika Sunan Kalijaga), Vols. Vol. 3, No. 1, no. ISSN : 2527-5836, p. 1– 13, 2018.

[13]. I. F. Rozi, E. N. Hamdana and M. B. I. Alfahmi, Pengembangan Aplikasi Analisis Sentimen Twitter Menggunakan Metode Naive Bayes Classifier (Studi Kasus SAMSAT Kota Malang), Jurnal Informatika Polinema, Volume 04, Edisi 02, no. ISSN: 2407-070X, 2018.

[14]. P. Antinasari, R. S. Perdana and M. A. Fauzi, Analisis Sentimen Tentang Opini Film Pada Dokumen Twitter Berbahasa Indonesia Menggunakan Naive Bayes Dengan Perbaikan Kata Tidak Baku, Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, Vol. 1, No. 12, no. e-ISSN: 2548-964X, pp. 1-9, December 2017.

[15]. A. F. Hidayatullah and A. SN, Analisis Sentimen dan Klasifikasi Kategori Terhadap Tokoh Publik Pada Twitter, Seminar Nasional Informatika 2014 (SemnasIF 2014) , no. SSN: 1979-2328, 2014.

[16]. M. R. Huq, A. Ali and A. Rahman, Sentiment Analysis on Twitter Data using KNN and SVM, (IJACSA) International Journal of Advanced Computer Science and Applications, vol. 8, 2017.

[17]. G. Chakraborty, M. Pagolu, S. Garla, Text Mining And Analysis Practical Methods, Examples, And Case Studies Using SAS, North Carolina, USA: SAS Institute Inc., 2013.

[18]. Balahur and R. Steinberger. 2009. Rethinking sentiment analysis in the news: from theory to practice and back. In Proceeding of the 1st Workshop on Opinion Mining and Sentiment Analysis (WOMSA).

[19]. Marco Baroni, Silvia Bernardini, Adriano Ferraresi, and Eros Zanchetta. 2009. The wacky wide web: A collection of very large linguistically processed webcrawled corpora. Language Resources and Evaluation, 43(3):209– 226.

[20]. John Blitzer, Mark Dredze, and Fernando Pereira. 2007. Biographies, Bollywood, Boomboxes and Blenders: Domain Adaptation for Sentiment Classification. In Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics (ACL), pages 440– 447.

[21]. Sabine Brants and Silvia Hansen. 2002. Developments in the tiger annotation scheme and their realization in the corpus. In Proceedings of the Third Conference on Language Resources and Evaluation, pages 1643–1649, Las Palmas.

[22]. Simon Clematide and Manfred Klenner. 2010. Evaluation and extension of a polarity lexicon for german. In Proceedings of the First Workshop on Computational Approaches to Subjectivity and Sentiment Analysis, pages 7– 13.

[23]. Hoa Trang Dang. 2009. Overview of the TAC 2008 Opinion Question Answering and Summarization Tasks. In Proceedings of the Text Analysis Conference (TAC), Gaithersburg, MD, USA.

[24]. Joseph L. Fleiss. 1981. Statistical Methods for Rates and Proportions. Wiley series in probability and mathematical statistics. John Wiley & Sons, New York, second edition.

[25]. J. Richard Landis and Gary G. Koch. 1977. The Measurement of Observer Agreement for Categorical Data. Biometrics, 33(1):159–174.

[26]. Bo Pang and Lillian Lee. 2005. Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales. In Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL'05), pages 115–124, Ann Arbor, Michigan, June. Association for Computational Linguistics.

[27]. Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. 2002. Thumbs up? sentiment classification using machine learning techniques. In Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing, pages 79–86. Association for Computational Linguistics, July.

[28]. Robert Remus and Christian Hanig. 2011. Towards Well- ¨ grounded Phrase-level Polarity Analysis. In Proceedings of the 12th International Conference on Intelligent Text Processing and Computational Linguistics (CICLing), number 6608 in LNCS, pages 380–392. Springer.

[29]. Yohei Seki, Lun-Wei Ku, Le Sun, Hsin-Hsi Chen, , and Noriko Kando. 2010. Overview of Multilingual Opinion Analysis Task at NTCIR-8 - A Step Toward Cross Lingual Opinion Analysis. In Proceedings of NTCIR-8 Workshop Meeting.

[30]. Veselin Stoyanov and Claire Cardie. 2011. Automatically creating general-purpose opinion summaries from text. In Proceedings of the International Conference Recent Advances in Natural Language Processing 2011, pages 202–209, Hissar, Bulgaria, September. RANLP 2011 Organising Committee.