

Sentiment Analysis in Facebook using Machine Learning Techniques

Naveen Kumar G¹, Nethravathi G², Nidhi Kulkarni³, Reshma T⁴, Roopa⁵
^{1, 2, 3, 4, 5}Ballari Institute of Technology and Management, Ballari, Karnataka, India.

Abstract: *Sentiment analysis is the process of computationally identifying and categorizing opinions expressed in text to determine the users' attitude towards a particular product is positive, negative or neutral. In today's world, opinions and reviews accessible to us are one of the most critical factors in formulating our views and influencing the success of a brand, product or service. With the advent and growth of social media in the world, stakeholders often take to expressing their opinions on popular social media, namely facebook. While Facebook data is extremely informative, it presents a challenge for analysis because of its humongous and disorganized nature. This paper is a thorough effort to dive into the novel domain of performing sentiment analysis about people's opinion about Google and apple productions. For performance analysis machine learning approaches such as n-gram technique is used. Hybrid approach is implemented by combining lexicon-based and machine learning algorithms. The result obtained through this approach is feasible to perform sentiment analysis in facebook with high accuracy.*

Keywords: *Sentiment, Facebook, N-gram, Lexicon-approach.*

I. INTRODUCTION

Social Network Sites (SNS) such as Facebook, Twitter, and online forum play important roles in our daily lives where people are allowed to communicate and share information and exchange opinions about celebrities, politicians, products and companies. Social Media has captured the attention of the entire world as it is thundering fast in sending thoughts across the globe, user friendly and free of cost requiring only a working internet connection. People are extensively using this platform to share their thoughts loud and clear. Facebook is one of the most popular online SNS. It contains comments posted by people where they can express their emotions and opinions via texts and emoticons. This paper proposes to quantify sentiments in facebook by analyzing the users emotions towards a product based on comments in a post. Sentiment analysis can be performed using machine learning, lexicon, NLP or hybrid techniques. There are many enhancement methods to enhance sentiment analysis results such as feature selection, data integration and crowd sourcing. Applying sentiment analysis over big data leads to a lot of insights and business benefits. Sentiment analysis, opinion mining or emotion detection is the process of extracting sentiment from text which is commonly used over online unstructured text like social media data streams. Data mining techniques are used for the extraction of logical patterns from the structured database. Text mining techniques are very complex than data mining due to unstructured and fuzzy nature of natural language text.

II. RELATED WORKS

The paper[1] titled "sentiment analysis in facebook and its application to e-learning" presents a new method for sentiment analysis in facebook that, starting from messages written by users, supports: (i) to extract information about the users' sentiment polarity(positive, neutral or negative), as transmitted in the messages they write; and (ii) to model the users' usual sentiment polarity and to detect significant emotional changes.

The paper[2] titled "sentiment analysis over social networks: an overview" presents that sentiment analysis can be applied in four levels: sentence, aspect and document and user level. this can be performed using machine learning (clustering or classification), lexicon, nlp, ontology or hybrid techniques. there are many enhancement methods to enhance sentiment analysis results such as feature selection, data integration, data cleaning, and crowd sourcing.

The paper[3] titled "Sentiment analysis of social networking sites (sns) data using machine learning approach for the measurement of depression" presents how to find the depression level of a person by observing and extracting emotions from the text, using emotion theories, machine learning techniques, and natural language processing techniques .In this paper a comparison is made among svm, nb and me classifiers regarding sentence level sentiment analysis for depression measurement. Our experiment indicates that svm shows superior result as compare to nave bayes and maximum entropy classifiers. we observed that the accuracy of svm is 91 %, the accuracy of nave base is 83 % and the accuracy of maximum entropy is 80 %.

The paper[4] titled “Sentiment analysis and opinion mining using machine learning techniques” presents that all human activities is based on opinions and opinions are key values of human operations. Analysis of sentiments is a multidiscipline area that covers NLP, text mining, and machine learning. Different machine learning algorithms such as random forest, naïve bayes, SVM are used which gives most appropriate accuracy.

III. DATA AND METHODOLOGY

A. Existing System

Initially in the existing system iteration process is done (for example 1st iteration they are going to take 100 comments, within that 100 comments which words are coming with more positive or more negative count that words will be added as positive or negative before 2nd iteration).Here accuracy is less because after iteration immediately we are considering positive or negative sentiment without considering left out words in comments.

1) Disadvantages of Existing System

- a) Time consuming
- b) Accuracy is less

B. Proposed System

We are going to propose the system for detecting sentiment for dynamic comments based on the threshold concept. Based on the threshold value we are going to achieve accuracy in the project. In the proposed system initial expansion is done based on the topic selected. Based on the topic every word in the particular comment, sentiment type of word is checked. Finally positive, negative or neutral count is incremented. The left out word which is not in any sentiment type ,that word sentiment is decided based on the positive ,negative and neutral count in that particular comment. If positive count is more than negative and neutral them it will be considered positive sentiment only. Finally if left out word sentiment crosses threshold value, then that word sentiment is permanently considered.

1) Advantages of the Proposed System

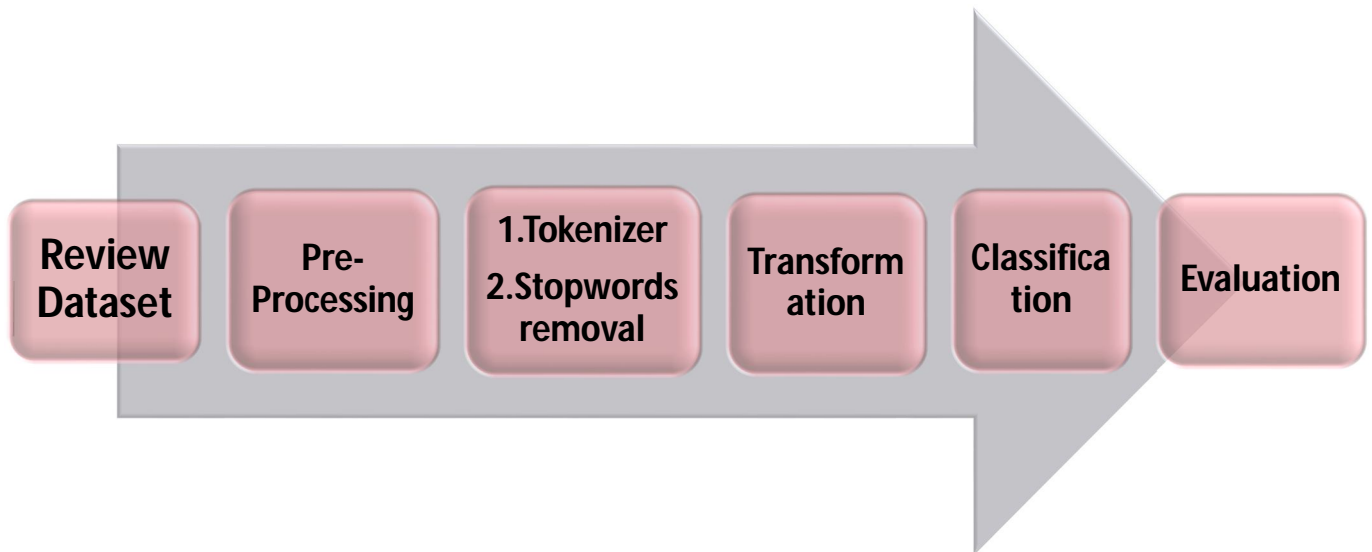
- a) Accuracy
- b) Efficiency

IV. IMPLEMENTATION DETAILS

The implementation stage in system project involves careful planning investigation of the current system and its constraints on implementation design of the methods to achieve change over etc. The errors in the code will be rectified during the phases of testing.

A. Major Modules

- 1) **Data Import Module:** In this module, comments are retrieved from the facebook API dynamically based on the product name input. To retrieve comments from the facebook API account, first need to create facebook account in developer’s console. After creating account we will get consumers token key and access token key, with the help of generated keys, we are going to communicate with facebook API to retrieve comments. The retrieved comments are imported into database.
- 2) **Preprocessing Module:** In this module, the comments which are imported to database from the facebook API, these comments consist of unnecessary words, whitespaces, hyperlinks and special characters. First we need to do filtering process by removing all unnecessary words, whitespaces, hyperlinks and special characters.
- 3) **Self Learning and word standardization System:** In this module, First we need to initialize the dictionary (first iteration dictionary).In the dictionary generally we need to initialize the positive, negative neutral and nouns. All big data and data mining projects based on the trained data, without trained data (initialization of words).So initialization of the trained data is very important. In the self learning system, we are doing word standardization ,here we are not considering past, present and future status of the words, only we are considering the word
- 4) **Sentiment Analysis Module:** In this module, preprocessed comments are fetched from the database one by one. First we need check one by one keyword whether that keyword is noun are not, if noun we will remove it from the particular comments. After that the remaining keywords checked with sentiment type, whether that keywords are positive sentiment or negative sentiment or neutral sentiment. The remaining keywords in the comment which does not belongs to any of the sentiment will be assigned temporary sentiment based on the more count of positive, negative and neutral. In the second iteration if the remaining word crosses the threshold of positive, negative or neutral, that keyword permanently added as expansion in the dictionary. Finally sentiment of the comment is detected based on the positive, negative and neutral words in the particular comment.



Flow Diagram

Fig: 4.1 Steps in Sentiment analysis process.

The above figure shows the steps in sentiment analysis process. First step is to collect the dataset (comments) from the Facebook. Collected data need to be pre-processed i.e., removal of unnecessary words, whitespaces, hyperlinks and special characters. Next step is we need to initialize a training data where positive, negative and neutral keywords are stored. Pre-processed data is checked word by word and if it matches any of the keywords in the trained data then it is classified as positive, negative or neutral. This process is done throughout the dataset and finally we will get the total count of positive, negative and neutral words present in the dataset. Later performance analysis of the process is done.

V. RESULTS AND CONCLUSION

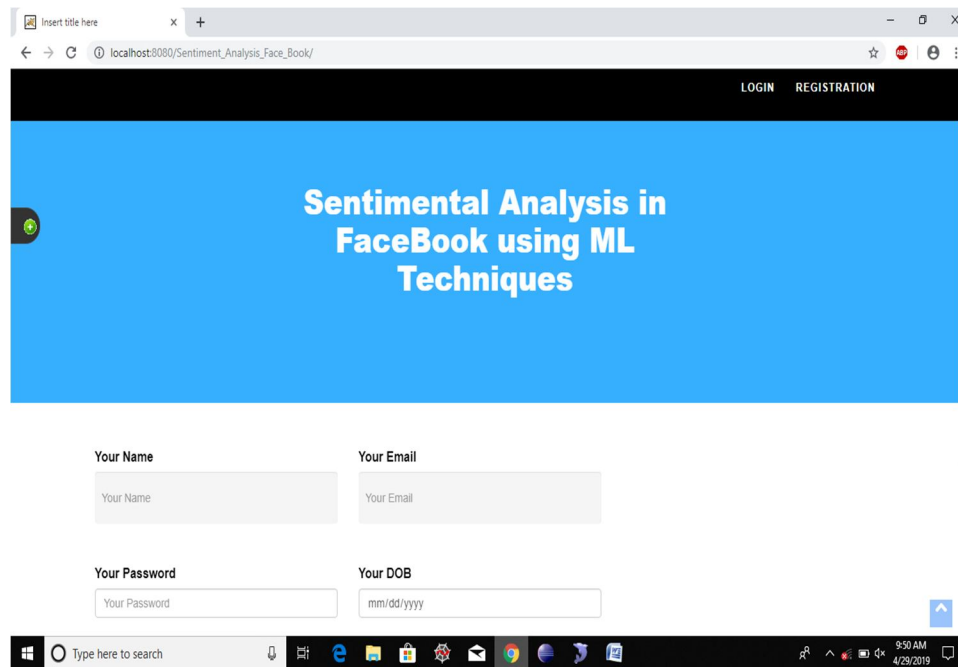


Fig 5.1: Registration Form

The above figure shows the Registration from where a user needs to register themselves by filling their details such as name, e-mail, and password.

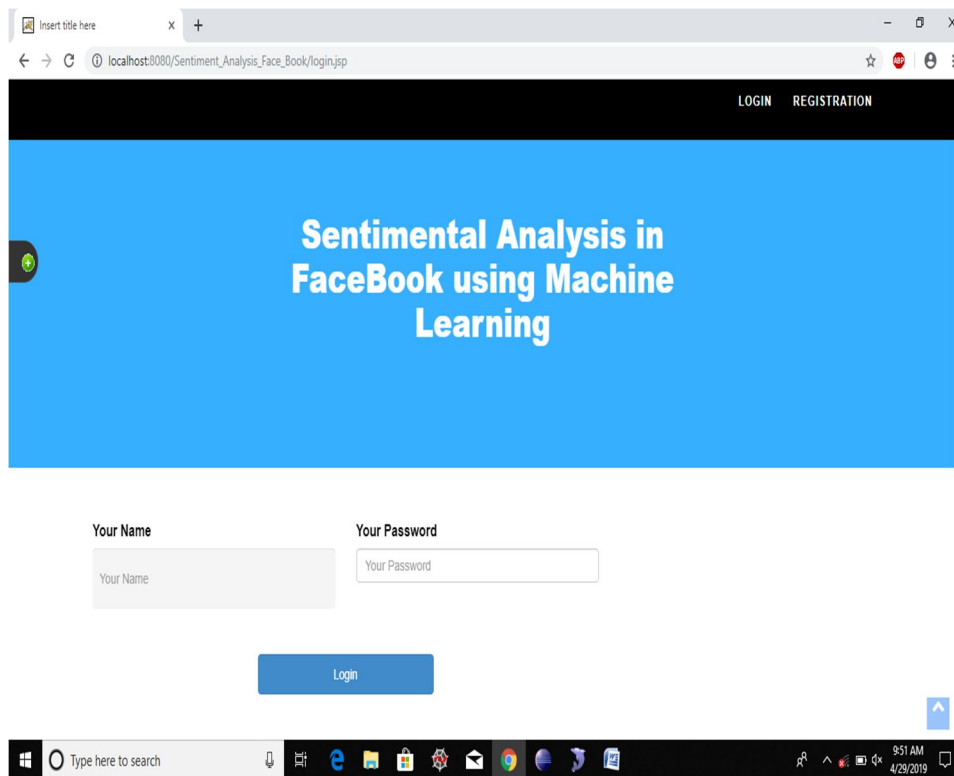


Fig: 5.2 Login Page

The above figure shows the login page where users need to enter his username and password. Authentication of the user is done by connecting to database to check whether he is a authorized user or not.

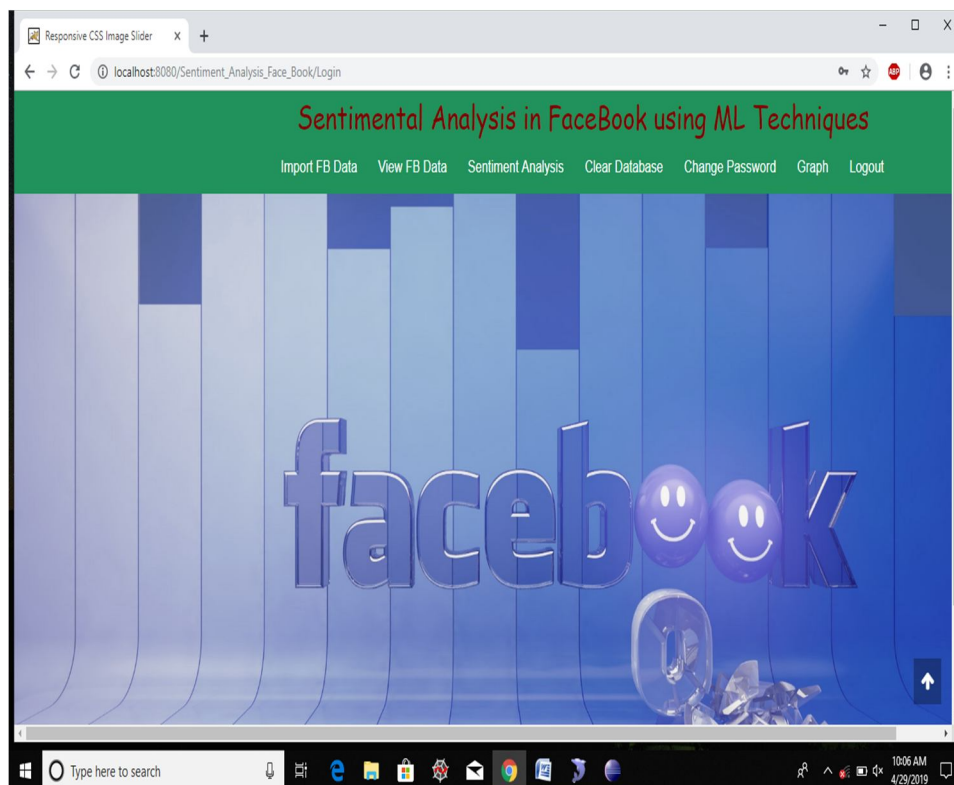


Fig: 5.3 Menu of overall process of sentiment analysis.

The above shows the menu where all the options required for the sentiment analysis process is shown.

- A. Loading data
- B. Viewing data
- C. Sentiment analysis
- D. Graph

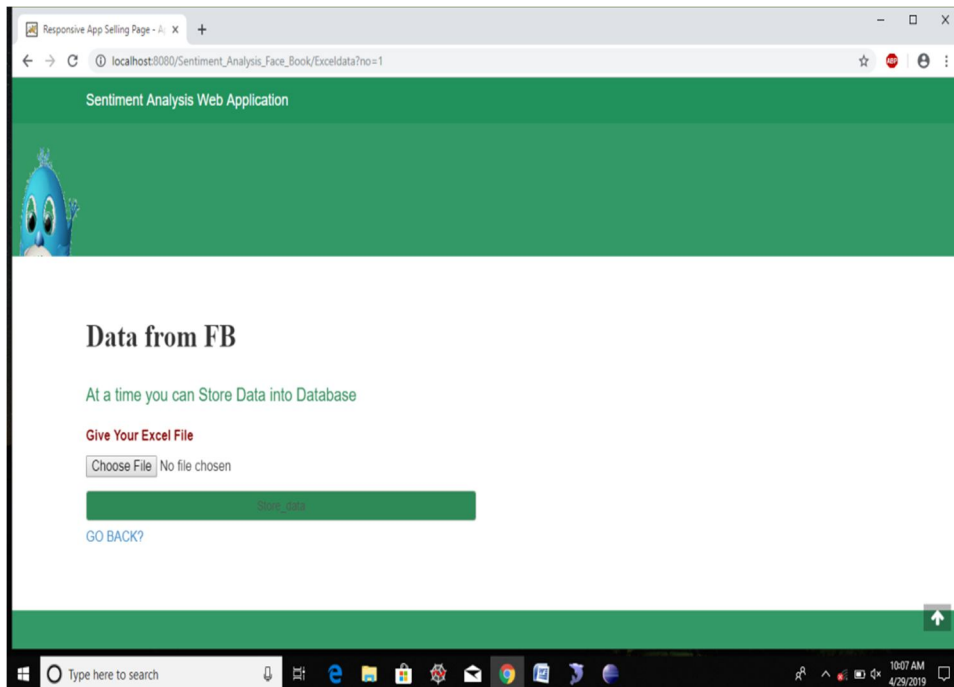


Fig: 5.4 Fetching dataset.

The above figure shows that we need to choose the dataset for the sentiment process.

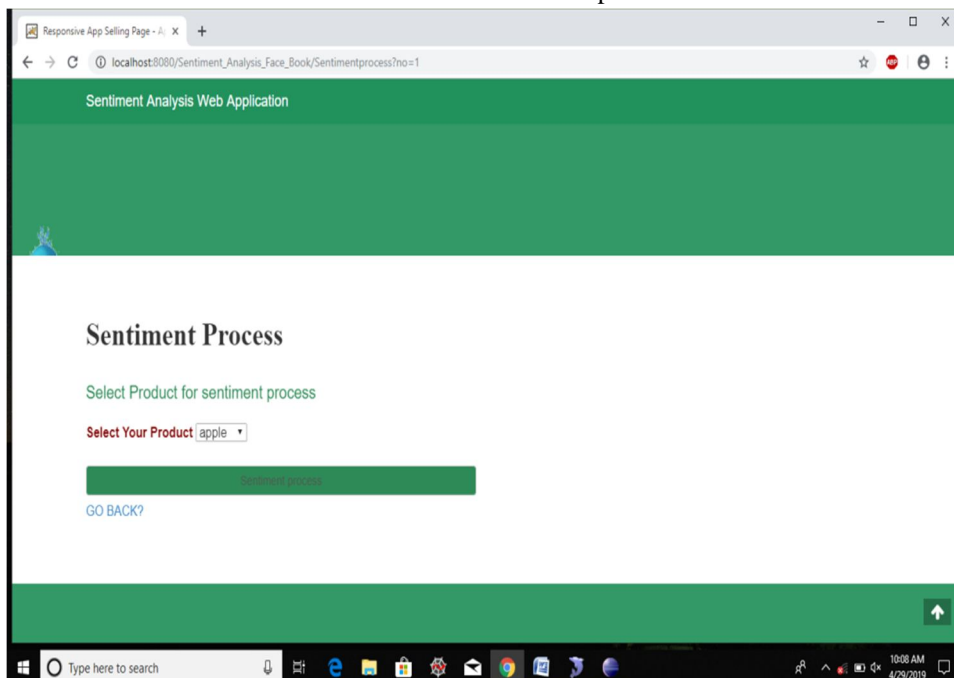


Fig: 5.5 Sentiment Process.

In above figure we need of select a product either Google or apple to know its polarity.

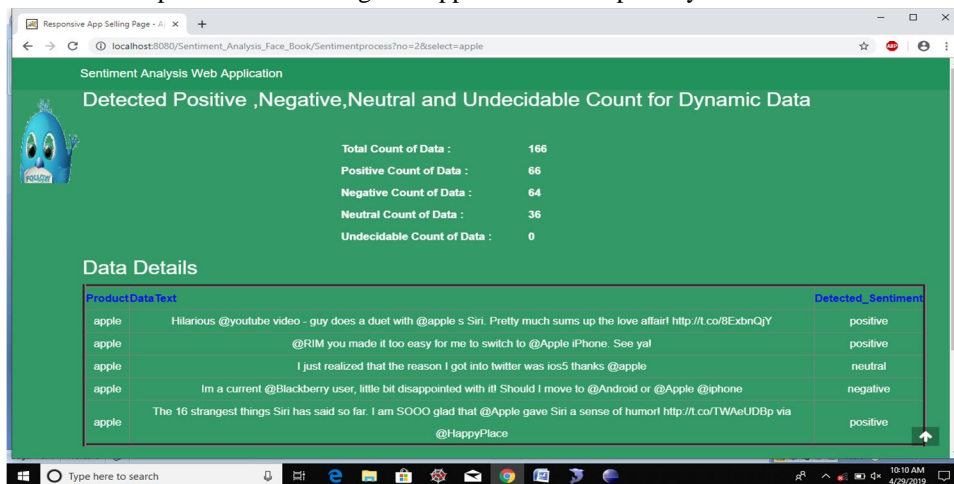


Fig: 5.6 Polarity Detection.

In the above sentiment process has been done and it shows total count of data, number of positive words, number of negative words, and number of neutral words. And it also classifies which are positive, negative and neutral comments.



Fig 5.7 Output of sentiment process represented in graph.

The output of previous figure is represented in form of graph where comparison between two products can be made easily.

VI. CONCLUSION

Sentiment analysis is an effective way of classifying the opinions formulated by people regarding any topic, service or product. Automation of this task makes it easier to deal with the massive amount of data being produced by social websites like Facebook on a real-time basis. Compared to the existing system, based on the threshold concept, iteration concept, keywords are expanded. Based on the expanded words efficiency of the system is increased.

REFERENCES

- [1] Alvaro Ortigosa, Jose M Martin, Rosa M Carro, "Sentiment Analysis in Facebook and its application to E-learning". Department of Computer Science, Universidad Autonoma de Madrid, Francisco Tomas Y Valiente 11, 28049 Madrid, Spain.
- [2] Khaled Ahmed, Neamat el Tazi, Ahmad Hany Hossny "Sentiment Analysis over Social Networks: An Overview" – 05 October 2015.
- [3] Federico Neri, Carlo Aliprandi, Federico Capeci, Montserrat Cuadros, "Sentiment Analysis on Social Media" 2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. -01 June 2014
- [4] Ramya Mala P, Seedhana Devi S, "Product Response Analytics In Facebook". Department Of Computer Science, Sri Vidya College Of Engineering And Technology, Virudhunagar. – ICICCS 2017.