**ABESEC Ghaziabad**



**Department of Computer Science & Engineering**

**SYNOPSIS REPORT**

**(Session 2022-23)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Project Title:**  **Opinion mining of public reviews using sentiment analysis system** | | | | |
| **Project Type**(application, product, research, review, etc.) | |  | | |
|  | **Name** | **Roll Number** | **Section** | **Signature** |
| **Group member (1)** | **Nitin Singh** | **2000320100109** | **B** |  |
| **Group member (2)** | **Karan Singh** | **2000320100078** | **B** |  |
|  |  |  |  |  |
| **Project Guide** | **Mr. Abhishek Kumar Shukla** | **Remarks:** | | |
| **Signature** |  |
| **Dateofsubmission** | **24-12-2022** |

**INDEX**

|  |  |  |
| --- | --- | --- |
| **S. No** | **Topic** | **Page. No** |
| **1** | **Introduction** | **2-3** |
| **2** | **Motivation** | **3-4** |
| **3** | **Objective** | **4-5** |
| **4** | **Scope of project** | **5-6** |
| **5** | **Related previous work** | **6-7** |
| **6** | **Diagram and Flow chart** | **7-8** |
| **7** | **Implementation and Results** | **8-9** |
| **8** | **Conclusion and future work** | **10-11** |
| **9** | **References** | **11** |

**INTRODUCTION**

**Problem introduction:**

The world of internet has brought us everything at ease, from buying products to searching the next new product purchases. When you have a platform to host a great website to promote the products which you have, the internet is now turning into a place where consumers check out products and explore things. Then they take decisions accordingly.

It is very important to imagine the value of customer opinions, but there is nothing to highlight beyond the data in terms of how reviews are used and how they affect the business. Statistical methods analyze how customers behave prior to and after using services or purchasing products, which can help them improve business development plans.

Reviews can not only have the capability to affect consumer decisions but also build up the company's image. Reviews have the potential to win customer feedback, and encourage people to engage with the company. Customer communication ultimately leads to improved business benefits. Reviews have utmost importance in our daily lives from e-commerce websites, social media platforms to political reviews.

Classification of reviews as positive and negative is the need of the hour. The machine learning algorithms and tools can provide information by analyzing product reviews automatically and categorizing it by positive, neutral, and negative. We need a review sentiment analyzer system which can accurately do this task for us.

This project will cater the need to classify the reviews in an efficient manner.

By using sentiment analysis to organize product reviews, we can understand our customer’s likes and dislikes for our product, can compare our product reviews with competitors, and can get the latest product details.

**Motivation**

**According to consumer’s vision:**

While making a decision it is very important that we know the opinion of the people around us. Earlier the group was usually small, with a few loyal friends and family members. But, now with the arrival of the Internet we see people expressing their opinions on blogs and forums.

These are now being studied diligently by people who want an idea about a particular business (product, movie, etc.).

Therefore, there are a lot of ideas available on the Internet. From a consumer perspective, getting ideas about a particular business is important. To try transferring such a large amount of data to understand the general view is impossible users with a large amount of this data. Therefore, there is a need for a system for positive reviews and negative reviews. In addition, writing these articles about their feelings will provide a brief summary of students with a general view of the business.

**According to manufacturer’s vision:**

With the explosion of Web 2.0 platforms like blogs, forums, etc., Consumers have a platform for sharing their product knowledge and their ideas, good or bad in any way product or service.

According to Pang and Lee (2008) these words of consumers can have a profound effect in building the opinions of other consumers and, ultimately, their product integrity, their purchasing decisions, and their product representative.

As consumers begin to use the power of the Internet to increase their horizons, it has become an explosion of review sites and blogs, where users can see the benefits of a product or service and errors. These ideas therefore shape the future of a product or service. Vendors need a system which can identify styles in customer reviews and use them to improve their product or service as well as identify future needs.

**According to community’s Vision:**

Recently, certain events, affecting the Government, have been caused by the use of the Internet. Social networks are used to bring people together to organize mass gatherings and to oppose oppression.

On the black side, social networks are used to suggest people about race or the human race, which has led to massive loss of life. Therefore, there is a need for Sentiment Analysis systems can identify such items and reduce them if necessary.

**Objective**

Sentiment analysis basically involves analyzing the information under a text and classifying the polarity of the opinion according to negative and positive parameter. In decision making, the preconditioning of others have a major impact on the convenience of customers for taking main decisions about online purchases. Sentiment is one of the many areas of computer studies that deal with natural language-based analysis. Such theoretical studies include, among other things, classification, emotional recognition and emotional impact, quality, value calculation, ideas in the text, identifying the source of the text and summarizing the concepts. Sentiment analysis has emerged as an exciting new trend in social media with a wide range of active applications ranging from business programs (intelligent marketing, benchmarking and benchmark performance and optimization), applications such as subcontracting technology.

Sentiment Analysis, the site of Natural Language Processing (NLP), is used to classify reviews using the sense of words to be classified as positive or negative. Using the sense expressed in words or text, ideas in any entity can be divided into positive or negative. For example, the phrase, ‘I am not happy with this product even though it is very cheap’ expresses negative feelings about this product. The level of feeling used is also considered For example, 'I like this product' shows a much better feeling than the sentence 'I like this product'. Apart from the common adjectives such as 'good', 'bad' and 'very good', conjunctions such as 'but', 'although', 'while' also have a voice in the full view of the sentence. Many challenges as organizations and individuals try to analyze and understand the opinions of others. Unfortunately to find sources of information monitor and analyze it are herculean activities. It is impossible to manually find sources online, extract ideas from them and express them in a common format. In recent years, millions of people express unresolved opinions about the various product features and their nuances. This creates an effective response that is important not only for product development companies, but also for competitors and other potential customers. Hence, the main objective of our project is to meet the above requirements and classify the reviews, sentiments or opinion as positive or negative. It uses neural networks and deep learning algorithms that learn on their own and do not require guidance and classifies the reviews accurately

**Scope of the Project**

In today’s scenario, social platforms are shooting up; the vast data can be used to meet business objectives, marketing, and other promotional strategies for their profits. The benefit of social media to mine public opinions and analyse their emotions can be obtained by opinion mining techniques.

Major scope of opinion mining include: -

• By using sentiment analysis different customer segments of your business analysis gets easy and helps us to have a better understanding of sentiment and opinions of people.

• Opinion mining can be used for analyzing political opinions.

• In stock market analysis.

• Movie recommendation systems.

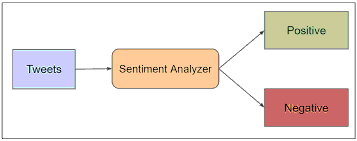
**Related Previous** **Work**

Sentiment Analysis and classification in the field of Machine learning can be performed in two ways. One method is supervised learning and another is unsupervised learning. Support vector machines (SVM) and naïve Bayes (NB) widely use monitoring techniques. Solutions associated with machine learning solutions involve the construction of separators from a document, where we can represent each text as a bag of words. Also, it is common to use certain methods to prevent and eliminate word loss. In general, the ethical categories in the domain in which they are trained do not reflect the same behavior in another domain because they rely heavily on the training data used.

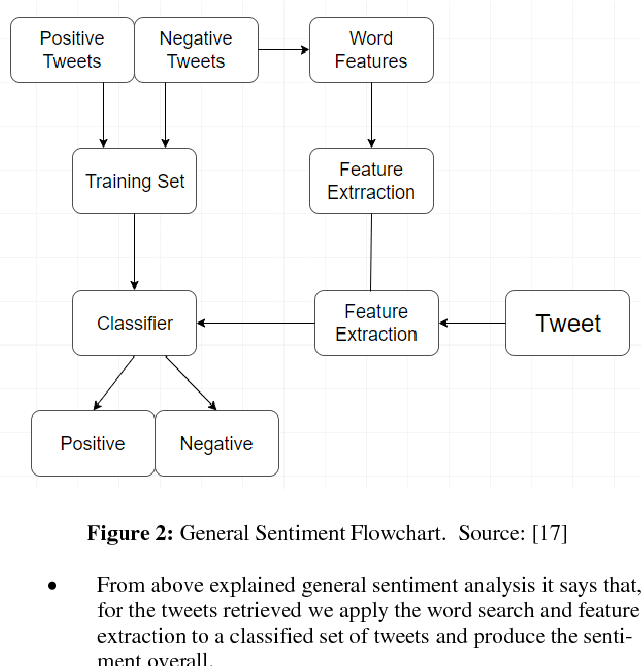
Many research papers use machine learning techniques along with LDA Analysis on naive bayes. Unsupervised machine learning technique is also used for this purpose. The machine learning algorithm is used to categorize the positive and negative reviews. Preprocessing is also done on the unstructured data. Building the vocabulary and extracting the features are a very important part of the process. This research work uses naïve bayes algorithm because of its high accuracy. Sentiment analysis can be used in a lot of decision-making areas also.

Sentiment Vader (VALENCE AWARE DICTIONARY AND SENTIMENT REASONER) is sentiment analysis instrument that is used to analyze social media data. Senti word net is used for natural language processing tool to calculate frequency and significance of word. As we know that sentiment analysis is one of the most tedious and difficult tasks in the application of natural language because people also strive to analyze emotions accurately and the process of extraction of the features and classification is not so efficient and the model does not get the desired accuracy. To overcome this, we have tried to implement a sentiment analysis model using Deep Neural Networks that performs better and give better accuracy and classifies almost every review correctly. The initial training of a deep learning model was extremely time-consuming and often required millions of data points until it began to learn on its own. But with deep learning the model gets trained and it performs far better than the machine learning model and gets a better efficiency and accuracy. Feature extraction gets very efficient and faster by using the deep neural networks.

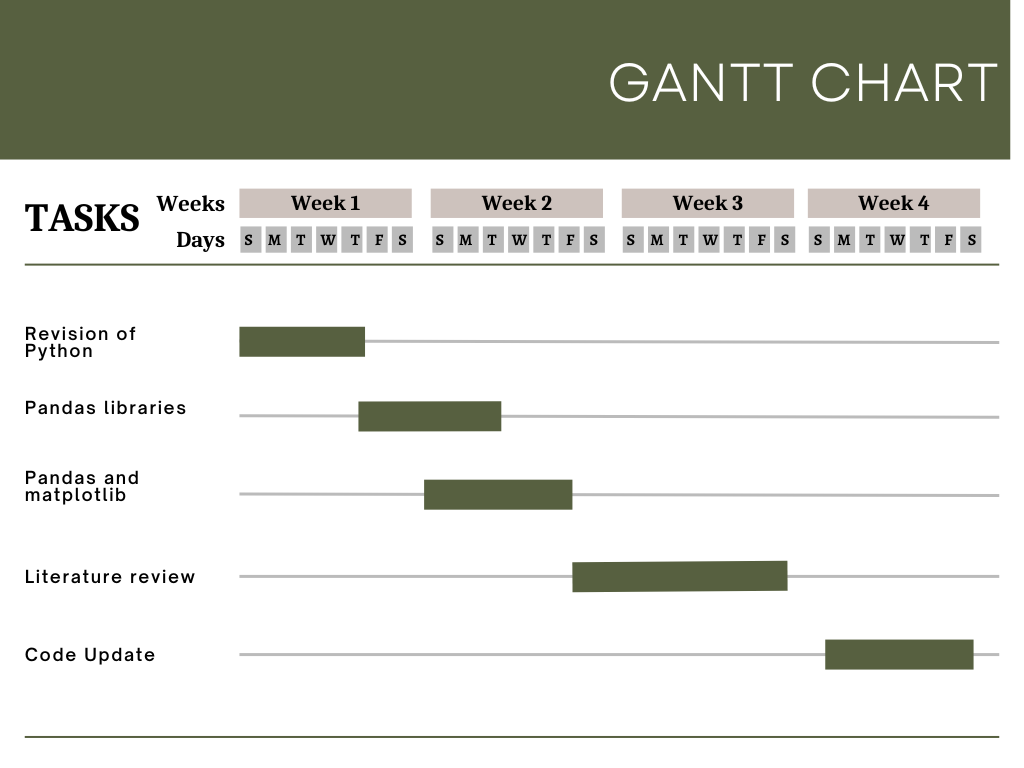
**Simple diagram**



**FLOW CHART**



**Gantt Chart**



**IMPLEMENTATION AND RESULTS**

**Software and Hardware Requirements**

**Software Requirements:**

**1. Python:** Python is an interpreter, high-level and general-purpose programming language. It was created by Guido van Rossum and it was first released in the year 1991. The language constructs and the object-oriented approach helped the programmers in order to write clear and logical code for the large-scale and small-scale projects.

**2. Jupyter Notebook:** The Jupyter notebook is basically an open-source web application which helps you to create as well as share the documents which contain life code visualization equations and a narrative text. It is also used for data cleaning as well as transformation, statistical modeling, data visualization, machine learning and much more.

**3. Google Colaboratory:** Google colaboratory is basically a cloud-based service which is used for the replication of the Jupiter notebook in the cloud. It does not require any installation on the system. Google colaboratory is basically used by those readers who use something other than a desktop to work through the examples.

**4. Python Libraries:**

**Matplotlib:** Matplotlib is basically a library for plotting in python language. It is used in order to provide an object-oriented application interface for including the plots into the applications. Most functions for plotting Matlab can easily be used in python. It includes different plots like bar plot, line plot, histogram, scatterplot etc. Using these types of plots, we can easily visualize the data.

**Pandas:** It is basically used for the cleaning of data and its analysis. Pandas provides various features such as exploring, transforming, visualizing as well as cleaning the data. It is basically an open-source python package. It is one of the most important tools for the data cleaning as well as analysis part.

**Keras :** Keras is a python-based framework which is assumed to be the coolest python library. It is known to easily represent the neural network problems. It is a modified version API of the tensorflow library. It is majorly used in data processing and in data visualization.

**Hardware Requirements:**

High CPU configuration, RAM minimum 8GB and processor 64-bit.

**CONCLUSION AND FUTURE WORK**

The trained model prepared classifies the reviews into negative and positive, basically identifying the sentiment associated with the review. Data is studied and analysed with the help of multi-layer perceptron model. Our model gives the result with the training accuracy of above 80 % and testing accuracy of 85% which is better than that comes with using machine learning algorithms and models like CNN and SNN. The model gives negligible overfitting, hence preferable to be used.

**References**

[1]Smija Dasthis and Mirsa Karim Sentiment analysis on textual reviews2018 *IOP Conf.*IOP Publishing Ltd *Ser.: Mater. Sci. Eng.* 396 012020​

[2] Anusuya dhara, SourishSengupta, Pranit Bose and Arladeb Saha Sentiment Analysis of Product-Based Reviews Using Machine Learning Approaches RCC INSTITUTE OF INFORMATION TECHNOLOGY 2017-18

​

[3] S.Muthukumara, A.Victoria ,Anand Mary Sentiment Analysis for Online Product Reviews using NLP Techniques and Statistical Methods International Journal of Mathematics And its Applications Volume 4, Issue 4 (2016), 303–312. ISSN: 2347-1557​

[4] MohdRidzwan and Yaakub, A Review on Sentiment Analysis Techniques and Applications  IOP Conf. Series: Materials Science and Engineering 2019 IOP Conf. Ser.: Mater. Sci. Eng. 551 012070​

[5] Brian Keith Norambuena∗ ,Exequiel Fuentes Lettura and Claudio MenesesVillegas,Sentiment analysis and opinion mining applied to scientific paper reviews, Universidad Catolica del Norte 2019

[6]AsadMasoodKhattak , RabiaBatool, Fahad Ahmed Satti, JamilHussain , Wajahat Ali Khan , AdilMehmood Khan and Bashir Hayat, Tweets Classification and Sentiment Analysis for Personalized Tweets Recommendation,2020 Hindawi Complexity Volume 2020, Article ID 8892552, 11 pages