Sentiment Analysis Using Natural Language Processing

Mr. Nikale Nikhil D, Mr. Patil Pratik V, Ms. Pawar Rutuja V, Ms. Phepale Rutuja B, Mr. Waghmare Suyash V

Department of Information Technology
Sanjivani Rural Education Society's
Sanjivani College of Engineering
Kopargaon - 423 603

Abstract: **Feeling examination can be viewed as significant utilization of AI, all the more especially NLP.As there are assortments of applications, Sentiment examination has acquired a ton of consideration furthermore is one among the quickest developing examination region in PC Science. It is a kind of information examination which is seen from news reports, client surveys, online media refreshes and so on Reactions are gathered and examined by specialists. All opinions can be grouped into three classes Positive, Negative also Neutral. The project gives a nitty gritty investigation of feeling investigation. It clarifies the sentiments of opinion examination, and furthermore various methodologies of opinion examination. The new instruments furthermore APIs alongside different certifiable uses of opinion investigation in different regions are likewise portrayed momentarily**

INTRODUCTION

Sentiment analysis is mining of text which identifies and extracts subjective information into source material, helping the business to understand the social sentiment of their brand, products or services while monitoring the online conversation. Use the sentiment analysis to determine the sentiments expressed in the user utterance. With the sentiment information you can manage the conversation flow or perform post-call analysis. For example, if the user sentiment is negative you can hand over the conversation to the actual human agent. Sentiment Analysis is a procedure used to determine if a piece of text is either Positive, Negative or Neutral. text analytics, natural language processing (NLP) and machine learning (ML) techniques used to determine whether sentiments are positive, negative or neutral. In text analytics, natural language processing NLP, machine learning techniques are combined to assign sentiment scores to the topics, categories or entities within the phrases. Sentiment analysis studies the subjective information in an expression, that is, the opinions, appraisals, emotions, or attitudes towards any topic, person or entity. Expressions can be classified into either positive, negative, or neutral. For example: "I really liked the functions of your device!" → Positive. Hybrid approach. Hybrid sentiment analysis models are most modern, efficient, and widely-used approach for the sentiment analysis. A social media sentiment analysis tells you about how all the people or users actually feel about your brand online. Rather than the simple count of mentions or comments, sentiment analysis considers emotions and opinions. It involves collecting and analyzing the information in the posts people share about your brand on social media.

MOTIVATION

This project aimed to create a system as it has always been a challenge for quantitative analysts and researchers. The Market depends on Sentiments and news. It is Help of various trends. Upcoming feature.

LITRATURE SURVEY

Sentiment analysis can be considered a major application of machine learning, natural language processing (NLP). As there are varieties of applications, Sentiment analysis is one among the fastest growing research area in computer science and has gained a lot of attention. Its a type of data analysis observed from user reviews, news reports, feedbacks, social media updates etc. Responses are collected and analyzed by the researchers. All sentiments can be classified into three categories i.e. Positive, Negative and Neutral. The paper gives the detailed study of the sentiment analysis. It explains the basics of sentiment analysis, its types, and different approaches of the sentiment analysis. The recent tools and APIs along with number of real world applications of sentiment analysis in the various areas are also described briefly.

```
In [51]: from sklearn.metrics import classification_report
         from sklearn.metrics import accuracy_score

         print(classification_report(X_test_targetSentiment, predictedGS_clf_linearSVC_pipe))
         print('Accuracy: {}'. format(accuracy_score(X_test_targetSentiment, predictedGS_clf_LinearSVC_pipe)))

                       precision    recall  f1-score   support

             Negative       0.93      0.95      0.94     19416
              Neutral       0.91      0.68      0.78      6353
             Positive       0.97      0.99      0.98     56999

             accuracy                           0.96     82768
            macro avg       0.94      0.87      0.90     82768
         weighted avg       0.95      0.96      0.95     82768

         Accuracy: 0.9553450608930988
```

```
In [39]: new_text = ["not really impressive. I had high hopes but I am Dissapointed"] #Negative
         X_train_targetSentiment[gs_clf_LinearSVC_pipe.predict(new_text)]
Out[39]: Negative    NaN
         Name: Sentiment, dtype: object

In [40]: new_text = ["I will say It is the one you should pick from the market"] #Positive
         X_train_targetSentiment[gs_clf_LinearSVC_pipe.predict(new_text)]
Out[40]: Positive    NaN
         Name: Sentiment, dtype: object

In [41]: new_text = ["does not works."] #Negative
         X_train_targetSentiment[gs_clf_LinearSVC_pipe.predict(new_text)]
Out[41]: Negative    NaN
         Name: Sentiment, dtype: object

In [42]: predictedGS_clf_LinearSVC_pipe = gs_clf_LinearSVC_pipe.predict(X_test)
         np.mean(predictedGS_clf_LinearSVC_pipe == X_test_TargetSentiment)
Out[42]: 0.9553450608930988

In [49]: from sklearn.ensemble import RandomForestClassifier
         clf_randomForest_pipe = Pipeline([("vect", CountVectorizer()), ("tfidf", TfidfTransformer()),
                                           ("clf_randomForest", RandomForestClassifier())])
         clf_randomForest_pipe.fit(X_train, X_train_targetSentiment)

         predictedRandomForest = clf_randomForest_pipe.predict(X_test)
         np.mean(predictedRandomForest == X_test_targetSentiment)
Out[49]: 0.9509109800089233
```

LIMITATION OF EXISTING SYSTEM

• Costing: The Existing system is high cost and this is main reason most of the system is failed.

• Technology Complexity: Most of system is the complex to understand, Not user friendly as compare to our proposed system

• Time Consuming Feature: In existing system, the performance is low and most of the time system gets hanged due to load.

• Not Easy to Understand: Systems re complex to understand and they were not user friendly

EXPERIMENTAL SETUP

1. Jupyter Notebook- The Jupyter Notebook is open-source web application that basically allows data scientists to create and share different documents that integrates the live code, equations, number of computational outputs, visualizations, and some other multimedia resources, along with the explanatory text in a single document.
2. Streamlit- Streamlit is open-source python framework for building web apps for Data Science and Machine Learning. We can instantly develop web apps and deploy them easily using Streamlit. Streamlit allows you to write an app the same way you write a python code.

Python is a high-level, general-purpose programming language. Its design philosophy emphasizes the code readability with the use of the significant indentation. The language constructs and the object-oriented approach aims to help programmers to write clear, logical code for small- as well as large-scale projects.[30] Python is dynamically-typed Programming Language and garbage-collected. It supports multiple programming paradigms, including structured , particularly procedural, object-oriented and functional programming. It is described as a "batteries included" language due to its comprehensive standard library.[31][32] It was Guido van Rossum who began working on Python in the late 1980s as a successor to the ABC programming language and first released it in 1991 as Python 0.9.0.[33]- Python 2.0 was released back in 2000 and introduced new features such as list comprehensions, cycle-detecting garbage collection, reference counting, and Unicode support. Python 3.0, released back in 2008, was a major revision that is not completely backward-compatible with earlier versions. Python 2 was discontinued with version 2.7.18 from 2020.

3. Back End- (Python) Python is an object-oriented, interpreted, high-level programming language with a dynamic semantic. Python is very simple, easy to learn syntax that emphasizes the readability of code and therefore it reduces the cost of program maintenance. Python supports modules, packages which encourages program modularity and code reuse.

SCOPE:

Project is implemented on the basis of data Security. The stake holders who participated in the requirement analysis process were the developers who will be among the end users of the security data for building block chain applications.

PROBLEM STATEMENT:

The sentiments are drawn from various texts and are classified accordingly into either positive, negative or neutral. sentiment analysis has the growing interest among the research community. To Analyze the sentiments based on dataset of reviews for products on E-commerce Platform (Amazon). Sentiment Analysis is a procedure used to determine if a Piece of text is positive, negative or neutral. In text analytics, natural language processing and machine learning techniques are combined to assign sentiment scores to the topics, categories or entities within the phrases.

SYSTEM ARCHITECTURE
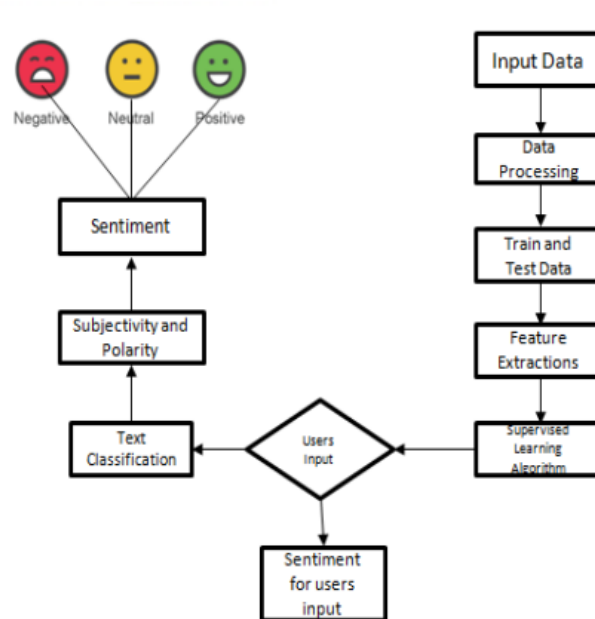
Fig -1: System Architecture Diagram

A System Architecture consist of multiple modules. The user will log in into the system by registering their details. As soon as the registration is completed, the system will store all their details in Database. Then the system will select the text which is observe in the review and the system will give/analyze the customer sentiment like positive, neutral, negative..

ADVANTAGES

• Data Processing: The system will process data to minimize the risk of error.
• Prediction: The system will the predict best result as per the input provided

METHODOLOGY

The single problem can be solved by different approaches/solutions. This actually considers the performance parameters for each approach. Thus it always considers the efficiency issues.

- Problem Solving Methods are concerned with the efficient realization of the functionality. This is an important characteristics of the Problem Solving Methods and should be deal with it explicitly.

- Problem Solving Methods achieve this efficiency by making assumptions on resources provided by their context (such as domain knowledge, etc) and by assumptions about the precise definition of the tasks. It is very important to make these assumptions explicit as it give the reason about Problem Solving Methods.

- The process of constructing the whole Problem Solving Methods is assumption-based. During this process assumptions are added that facilitate efficient open rationalization of the desired functionality

Result
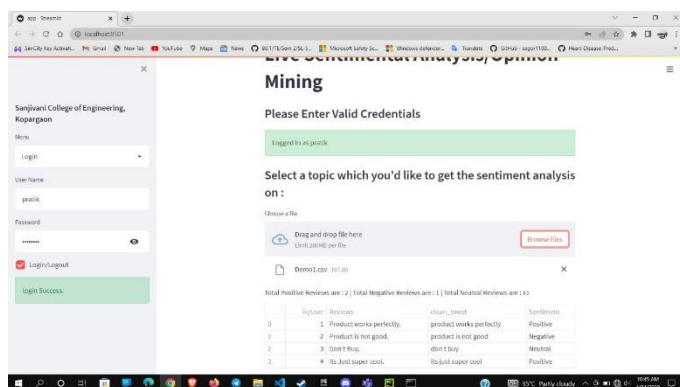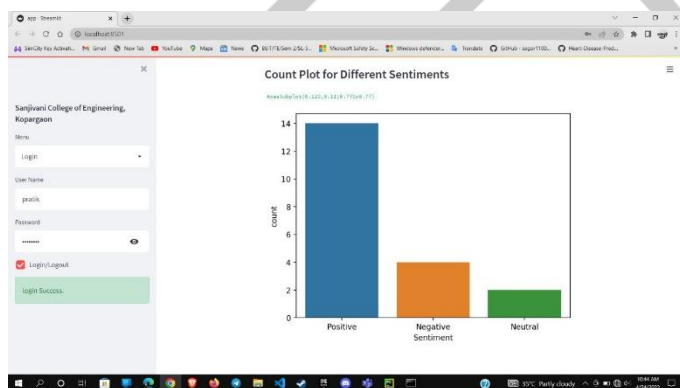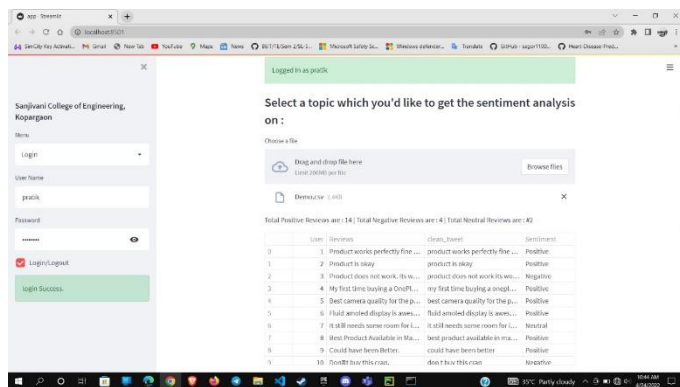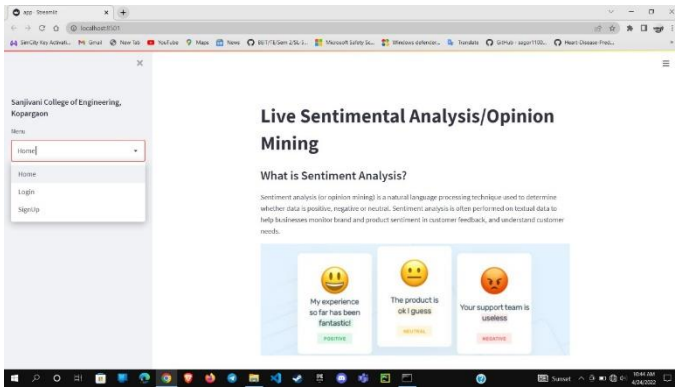
Machine Learning Model	Accuracy
Multinomial Naïve Bayes	84.45
Logistic Regression	88.78
Support Vector Machines	89.96
Decision Tree Classifier	92.91
Random Forest Classifier	95.09

5. CONCLUSION

Sentiment analysis is a field of study that analyzes people's sentiments, attitudes, or emotions towards any certain entities. This project tackles a fundamental problem of sentiment analysis, sentiment polarity categorization like positive, neutral and negative

REFERENCES

[1] Ms. Binju Saju, Ms. Siji Jose, Mr. Amal Antony,"Comprehensive Study on Sentiment Analysis: Types, Approaches, Recent Applications, Tools and APIs" IEEE 2020 Advanced Computing and Communication Technologies for High Performance Applications (ACCTHPA), 2020

[2] Tanjim Ul Haque, Nudrat Nawal Saber, "Sentiment Analysis on Large Scale Amazon Product Reviews", IEEE International Conference on Innovative Research and Development, 2018

[3] Shivangi Srivastava, Aastha Nagpal, Aashish Bagwari, "Various Approaches in Sentiment Analysis", 12th International Conference on Computational Intelligence and Communication Networks IEEE, 2020

[4] Pankaj, Prashant Pandey, Muskan, Nitasha Soni, "Sentiment Analysis on Customer Feedback Data: Amazon Product Reviews", International Conference on Machine Learning, Big Data, Cloud and Parallel Computing, 2018

[5] S.Padmaja et al, "Opinion Mining and Sentiment Analysis – An Assessment of Peoples' Belief: A Survey", International Journal of Ad hoc, Sensor Ubiquitous Computing IJASUC, Vol 4 No 1, Feb 2013.

[6] Tirath Prasad Sahu and Sanjeev Ahuja, "Sentiment Analysis of Movie Reviews: A study on Feature Selection Classification Algorithms", IEEE, 2016.

[7] Samha,Xu,Xia, Wong Li "Opinion Annotation in Online Chinese Product Reviews", In Proceedings of LREC conference, 2008.

[8] Nina Isabel Holleschovsky, "The social influence factor: Impact of online product review characteristics on consumer purchasing decisions", 5 th IBA Bachelor Thesis Conference, Enschede, The Netherlands, 2015