TWITTER SENTIMENT ANALYSIS AND NEWS ARTICLE PREDICTION

A Project Report submitted in partial fulfilment of the requirements for the awardof the degree of

BACHELOR OF TECHNOLOGY IN

COMPUTER SCIENCE AND ENGINEERING

Submitted by

KADIYALA P B ROHIT BHARADWAJ, 121910313006
SAMUDRALA MUNI VISHNU, 121910313011
SAI HARIDEEP, 121910313024
NAVEEN CHOWDARY, 121910313046

Under the esteemed guidance of

Dr. N SURESH KUMAR

Assistant Professor

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

GITAM

(Deemed to be University)

VISAKHAPATNAM

OCTOBER 2022

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING GITAM SCHOOL OF TECHNOLOGY

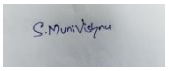
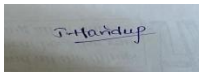
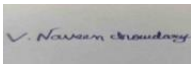
GITAM

(Deemed to be University)

DECLARATION

I/We, hereby declare that the project report entitled **"TWITTER SENTIMENT ANALYSIS AND NEWS ARTICLE PREDICTION**" is an original work done in the Department of Computer Science and Engineering, GITAM School of Technology, GITAM (Deemed to be University) submitted in partial fulfillment of the requirements for the award of the degree of B.Tech. in Computer Science and Engineering. The work has not been submitted to any other college or University for the award of any degree or diploma.

Date:

Registration No(s).	Name(s)	Signature(s)
121910313006	**KADIYALA P B ROHIT BHARADWAJ**	
121910313011	**SAMUDRALA MUNI VISHNU**	
121910313024	**SAI HARIDEEP**	
121910313046	**NAVEEN CHOWDARY**	

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING GITAM SCHOOL OF TECHNOLOGY

GITAM

(Deemed to be University)

CERTIFICATE

This is to certify that the project report entitled "TWITTER SENTIMENT ANALYSIS AND NEWS ARTICLE PREDICTION" is a bonafide record of work carried out by **KADIYALA P B ROHIT BHARADWAJ (121910313006), SAMUDRALA MUNI VISHNU (121910313011), SAI HARIDEEP (121910313024), NAVEEN CHOWDARY (121910313046**) students submitted in partial fulfilment of requirement for the award of degree of Bachelors of Technology in Computer Science and Engineering.

Project Guide	**Head of the Department**
Dr. N Suresh Kumar	**Dr. R.Sireesha**
Assistant Professor	**Professor**

TABLE OF CONTENTS

ABSTRACT

For many people, social media is fast expanding as a virtual communication medium. A great deal of information is being sent from user to user over the internet. Twitter is the most popular platform for social networking followed by WhatsApp. The sentiment of a given tweet is crucial in understanding the emotion of the tweet. We may readily comprehend the sentiments in tiny tweets. However, finding the sentiment in tweets with intricate phrases is more difficult.

This software collects all tweets with a certain keyword and produces the sentiment of tweets with that term. This allows you to determine whether a keyword has a positive, negative, or neutral response throughout the platform.

Everyone refers to news in their spare time. Everyone's life revolves around the news. When reading the news, one should be aware of the issue they are reading about, such as sports, crime, education, and so on. When a text is provided, this project identifies the topic of the text. It also anticipates the parts of speech in the text, providing us a fundamental understanding of how the English language works.

1. INTRODUCTION

In today's quickly changing world, many people communicate extensively using social media platforms, one of which is "Twitter." Many people connect and engage on Twitter by posting various information in the form of tweets. These might be Positive, Negative, or even Neutral. Based on a single term and a small number of tweets, this project attempts to analyze those tweets and forecast whether they will be favorable, negative, or neutral. The Twitter API is a crucial aspect of our sentiment analysis assignment since it allows us to retrieve recent tweets based on keywords and tweet count directly by connecting using API credentials. The tweets will be analyzed using several NLTK algorithms and approaches, such as VANDER, which generates sentiment intensity scores for each tweet. Finally, this project includes graph graphics to help you grasp the insights. Matplotlib was used to create this graphic.

There is a lot of news in the form of articles floating around the internet. The news can be divided into several categories that are difficult to distinguish. As a result, in this project, we also used some machine learning techniques to train and predict the news article's category based on previous training data, and a frontend is created using Streamlit for the articles NLP purpose, which helps to identify different parts of speech using tokenization, lemmatization, and word cloud.

2. LITERATURE REVIEW

[1] In Natural Language Processing Research, Sentiment analysis is a growing area of research with studies that mainly includes document level classification. In information gathering habit the key aspect has always been is finding out what other people think. The increased availability of personal blogs and online review sites which are major opinion-rich resources that have resulted in new possibilities and difficulties as majority of individuals `utilise such information technology to understand and comprehend other' perspectives. The sudden increased activity in the field of sentiment analysis and opinion mining that deals with the sentiment, computational opinion treatment, and test subjectivity, has thus occurred due to response to the increased interest in new systems that are dealing with opinions as a first-class object. (**Pang and Lee 2008**)

[2] We find and validate limitations from conjunctions on the positive or negative semantic orientation of conjoined adjectives using a large corpus. When each conjunction is analysed individually, a log-linear regression model achieves 82% accuracy in predicting whether conjoined adjectives are of the same or different orientations. A clustering method splits the adjectives into groups of distinct orientations by combining the restrictions across multiple adjectives, and then the adjectives are labelled positive or negative. Evaluations on real data and simulated trials show that the classification precision is greater than 90% for adjectives found in a small number of conjunctions in the corpus. (**Hatzivassiloglou and McKeown 1997;**).

[3] SentiWordNet is a WordNet-derived opinion lexicon in which each phrase is connected with numerical scores expressing negative and positive sentiment information. Opinion mining (OM) is a new sub discipline at the intersection of computational linguistics and information retrieval that is mostly deals with the opinion expressed in a text rather than the topic of the document. OM has a wide range of applications, from tracking user views ranging from political candidates to products stated in internet forums and customer relationship management. Recent research has attempted to automatically assess the PNpolarity of subjective phrases, i.e. if a term that

is of an opinionated material has a negative or positive meaning, in order to assist in the gathering of views from the text. SENTIWORDNET is free to use for research purposes and has a Web-based graphical user interface. **(Esuli and Sebastiani 2006)**

[4] The richest languages linguistically such as Azerbaijani have received little attention in the field of language natural processing. The text corpora developed from Azerbaijani news stories is intended to be used for automated news labelling with supervised machine learning algorithms. Chi-squared test and LASSO techniques were used for pre-processing and feature selection. Supervised machine learning approaches to the text corpus were applied to compare the performance results of well-established supervised machine learning algorithms in the area of Azerbaijani language (U Suleymanov and S Rustamov)

3. PROBLEM IDENTIFICATION & OBJECTIVES

Twitter is a massive social media network that is now available. Many individuals use Twitter to discuss or communicate about the concerns. It is critical to comprehend what the author of the tweet is stating, whether it is favourable, negative, or neutral. Small keywords are easy to grasp, while complicated ones make it more difficult to understand the emotion of the tweet. This project assists in quickly determining the sentiment of all tweets through the use of visuals for improved comprehension.

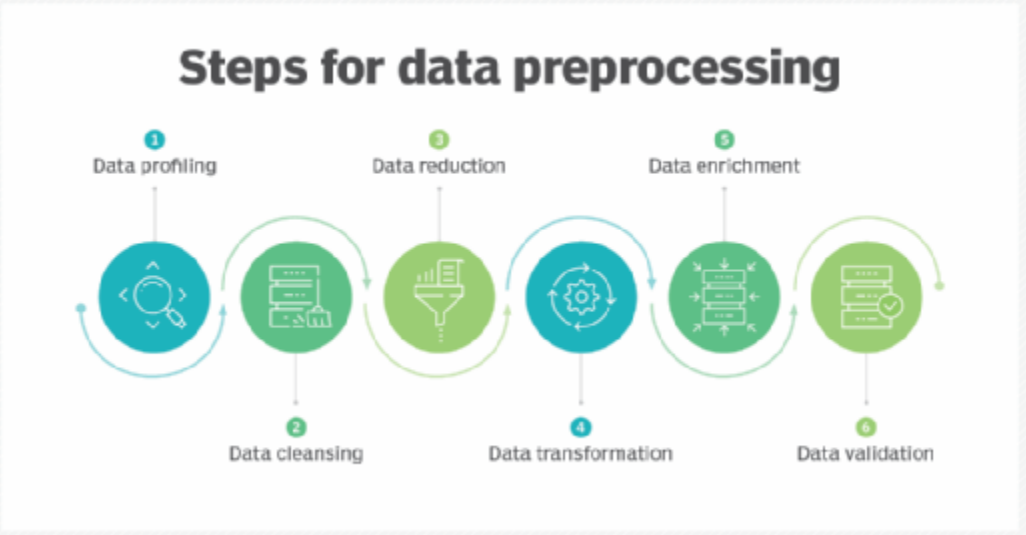
There is a lot of news flying around the internet in the form of articles. The news may be classified into various difficult-to-distinguish categories. One should be aware of the issue news to which they are referring. This is a really challenging undertaking. This project solves the challenges and forecasts the key categories of news such as sports, education, and so on, which may be extremely beneficial in quickly distinguishing the articles.

4. SYSTEM METHODOLOGY

4.1 DATA PREPROCESSING

The main objective of data pre-processing is preparing raw data and fitting it to a machine learning model. It is the first and most important stage in developing a machine learning model.

It is very difficult always to come across clean and prepared data when we work on a machine learning project. And before you can do anything with data, it must be cleaned and formatted. So we employ the data pre-processing job for this.



- Accomplished data pre-processing task in the second phase of project which is the prediction and categorization of news articles
- The data set was obtained from Kaggle, an open source platform for large datasets.

4.2 DESIGN:

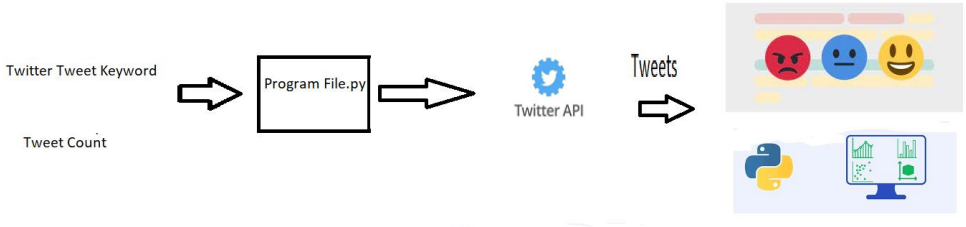


Figure a: Block Diagram, which describes working of Twitter tweet sentiment analysis

The above diagram can be divided into two parts. One is gathering the required tweets based on tweet keyword and count from the Twitter API.Next part is to analyse the gathered tweets by using NLTK sentiment analysis tools and provide a visualization output of overall tweets.

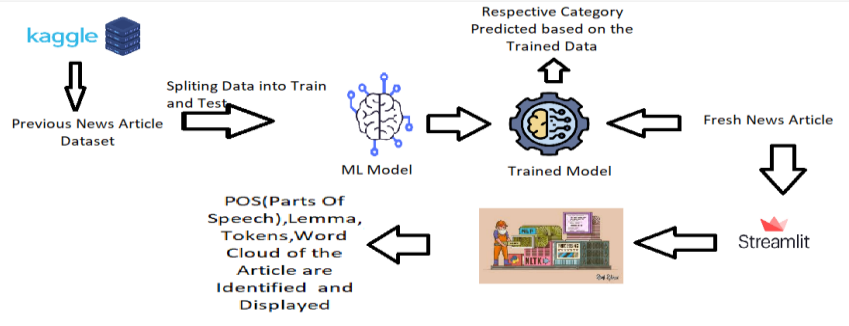
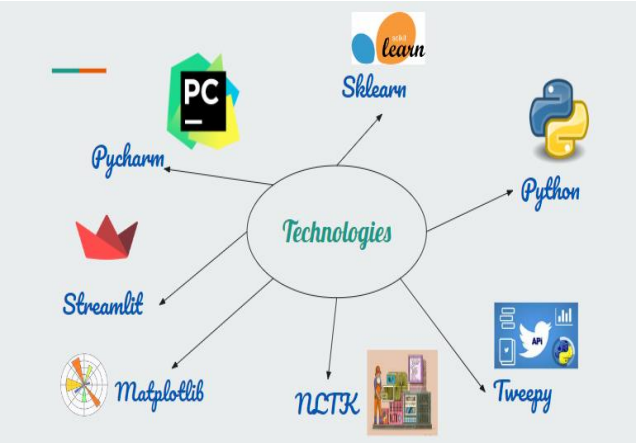


Figure b: Block Diagram, which describes working of News Article Prediction with ML and Its NLP outcome with a Frontend (Streamlit)

The method is separated into five steps, as shown in the above diagram. First, data from the News Article Category is acquired from the Kaggle Platform and converted into structured data using data pre-processing. The data is then divided into test and train data and trained with a Machine Learning Model. Once trained, the model will be able to recognise new news item categories based on the previously learned data set. We will receive a correctly predicted news category whenever we supply new input to the software. A front application is also constructed with Streamlit, which uses NLTK to detect the POS, Lemma, and Tokens, as well as show the article as a Word cloud.

5. OVERVIEW OF THE TECHNOLOGIES



5.1 TECHNOLOGIES USED:

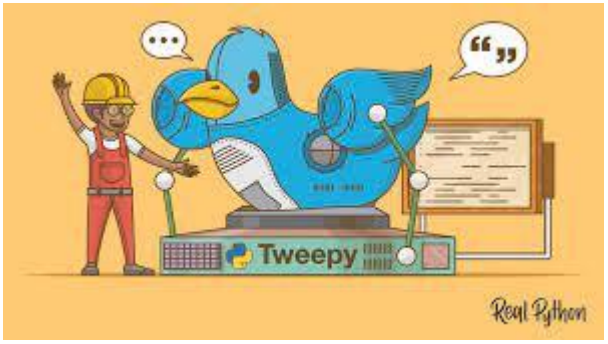
- Tweepy
- Twitter Developer API
- NLTK
- Streamlit
- Kaggle
- Scikit-learn
- Pandas
- Matplotlib
- Text Blob

5.1.1 NATURAL LANGUAGE PROCESSING:

It is a branch of linguistics, computer science, information technology, and artificial intelligence concerned with computer interaction with human (natural) language, specifically processing and analysing massive volumes of natural language data. It deals with computer programming. Natural language processing issues include speech recognition, natural language interpretation, and natural language generation. This aided us in determining the nature of the tweets based on the user supplied keyword and tweet count. It also helped us identify the POS (Parts of Speech), tokens, and lemma of words in news articles.

5.1.2 Tweepy

Tweepy Python programme available as an open source that provides a very straightforward method to use Python to access the Twitter API. Tweepy contains a collection of classes and methods that reflect Twitter's models and API endpoints, and it handles numerous implementation details transparently, such as data decoding and encoding. It is a simple Python module for interacting with the Twitter API. The API class provides access to all RESTful API services provided by Twitter. Each method can accept and return various parameters.



5.1.3 Twitter Developer API

Developer Platform

The Twitter developer site includes a suite of self-service tools for managing access to the Twitter API and Twitter Ads API.

You may do the following in the portal:

- Make and manage Twitter Projects and Apps (and the authentication keys and tokens that they provide).
- With the Twitter API premium v1.1 and v2 endpoints, you can manage your access levels and integrations.
- Learn more about the many endpoints and features that are available.
- If you have elevated access and an organisation account type, you can visit team pages where you may add and manage the many handles that have access to your team's account.

The Twitter API provides unique and sophisticated programmatic access to Twitter. We may also make use of major Twitter features such as Tweets, Direct Messages, Spaces, Lists, and People.

5.1.4 NLTK

The Natural Language Toolkit (NLTK) are set of applications and tools of Python for statistical and symbolic natural language processing in English. Various word processing products include a plethora of test datasets. You can use NLTK to perform a number of tasks, including: B. Visualization and tokenization of parse trees. The Natural Language Toolkit (NLTK) is a Python framework for developing algorithms that interface with human language data in statistical natural language processing applications (NLP). Text processing technologies include tokenization, parsing, classification,

and stemming, tagging, and semantic inference. The word tokenizes () function from NLTK is used to break a sentence into tokens or words. Tokens may be extracted from a string of characters using the tokenize.word tokenize()

5.1.5 Streamlit

reamlit is Python built open source framework app. It allows us to quickly construct data science and machine learning web apps. It works with popular Python libraries including scikit-learn, Keras, PyTorch, SymPy (latex), NumPy, pandas, and Matplotlib.

Streamlit, in a nutshell, convert data scripts into shareable web apps in minutes. Streamlit widgets enable you to incorporate interaction right into your apps, including buttons, sliders, text inputs, and more.

You may use Streamlit to work with any type of data or input. You can install Streamlit directly in our Python IDE by using the command pip install streamlit.

5.1.6 Kaggle

Kaggle is a data scientist and machine learning enthusiast community platform. Users can collaborate with others, publish and discover datasets, use GPU-integrated notebooks, compete and solve data science challenges on Kaggle with other data scientists. It is an open source platform for accessing different types of datasets.

5.1.7 Scikit-learn

Scikit-learn are perhaps the most helpful Python machine learning package. The sklearn package contains several powerful statistical modelling and machine learning techniques such as classification, clustering, regression, and dimensionality reduction. It provides a variety of effective statistical modelling and machine learning tools such as classification, clustering,

regression and dimensionality reduction through Python's consistency interface. Scikit-learn have a tonne of functionality available. Consider the following to have a better understanding of the spread:

• Supervised learning methods

• Cross-validation

• Unsupervised learning algorithms

• Various toy datasets

5.1.8 Pandas

Pandas are a data analysis and manipulation software package created for the Python programming language. It includes specific data structures and procedures for working with time series and mathematical tables. Wes McKinney came up with the moniker "Pandas" in 2008, and it refers to both "Panel Data" and "Python Data Analysis."

- With the aid of Pandas, we can examine large data sets and draw conclusions based on statistical principles.
- Pandas can organise disorganised data sets, making them understandable and useful.
- In data science and ML, relevant data is crucial.

5.1.9 Matplotlib

Matplotlib is a graphing library for Python programming language and NumPy, a numerical mathematics extension. For integrating charts into programmes utilising all-purpose GUI toolkits like Tkinter, wxPython, Qt, or GTK, it offers an object-oriented API. Matplotlib makes difficult things possible and simple things easy. Produce plots fit for publishing. Make interactive charts with zoom, pan, and update capabilities. It was created by John D. Hunter.

The majority of Matplotlib's tools are found in the pyplot submodule and are often imported using the plt alias. Matplotlib contains all types of plots which are used a lot in ML and Data science field. It is best suitable Library for Data Visualization

5.1.10 TextBlob

Being a Lexicon-based sentiment analyser, TextBlob It contains certain predetermined guidelines, or perhaps we should say a word and weight dictionary, which has some scores that assist in determining the polarity of a statement. It provides a simple API for getting started with common NLP operations such as tagging part-of-speech, extraction of noun phrase, sentiment analysis, translation, classification, and others.

For this reason, rule-based sentiment analysers are another name for lexicon-based sentiment analysers. Text blob acts as a Competitor for VANDER based Sentiment analysis which provides much more efficiency than the former one.

The vocabulary of a person, a language, or a field of study is called a lexicon.

6. IMPLEMENTATION

6.1 Coding

6.1.1 For Twitter Sentiment Analysis

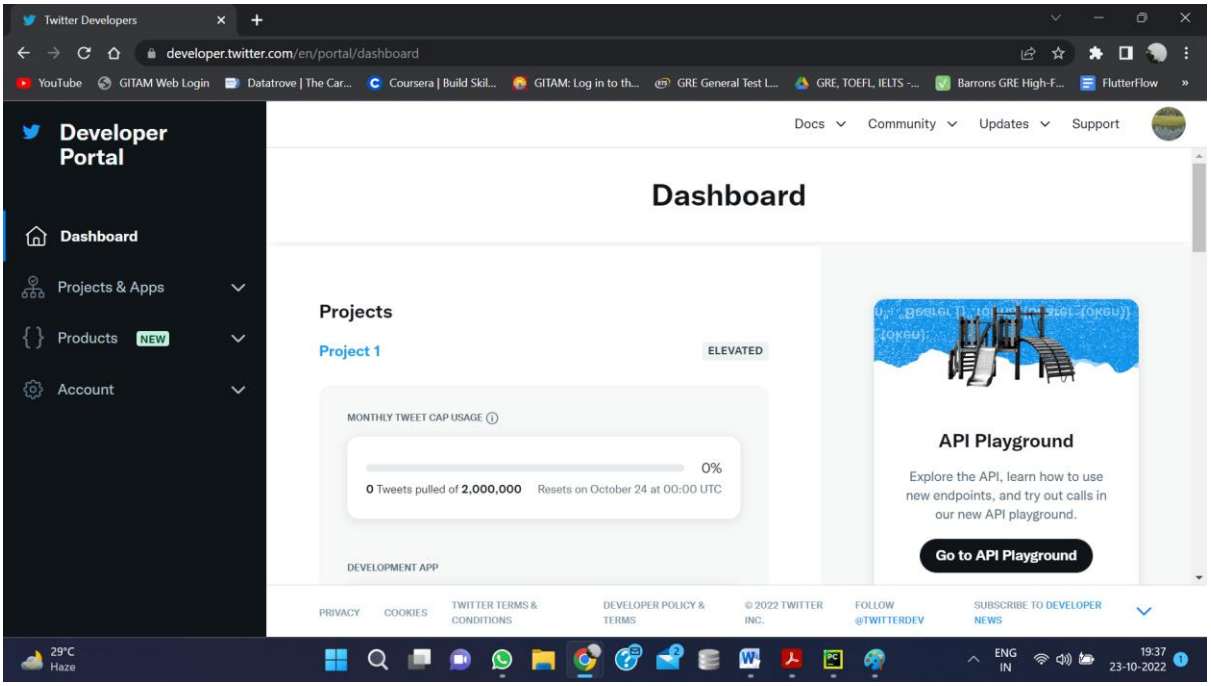


Fig 6.1.1-a: Twitter Developer API

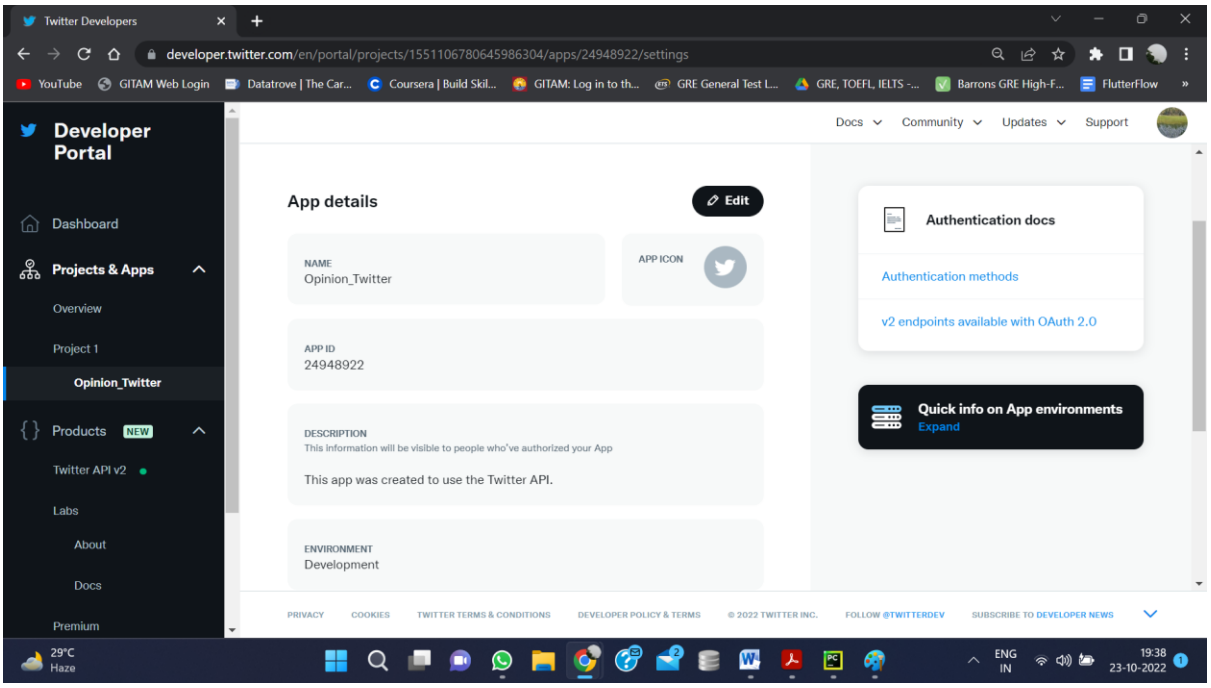


Fig 6.1.1-b: Twitter Developer API Project Creation Details

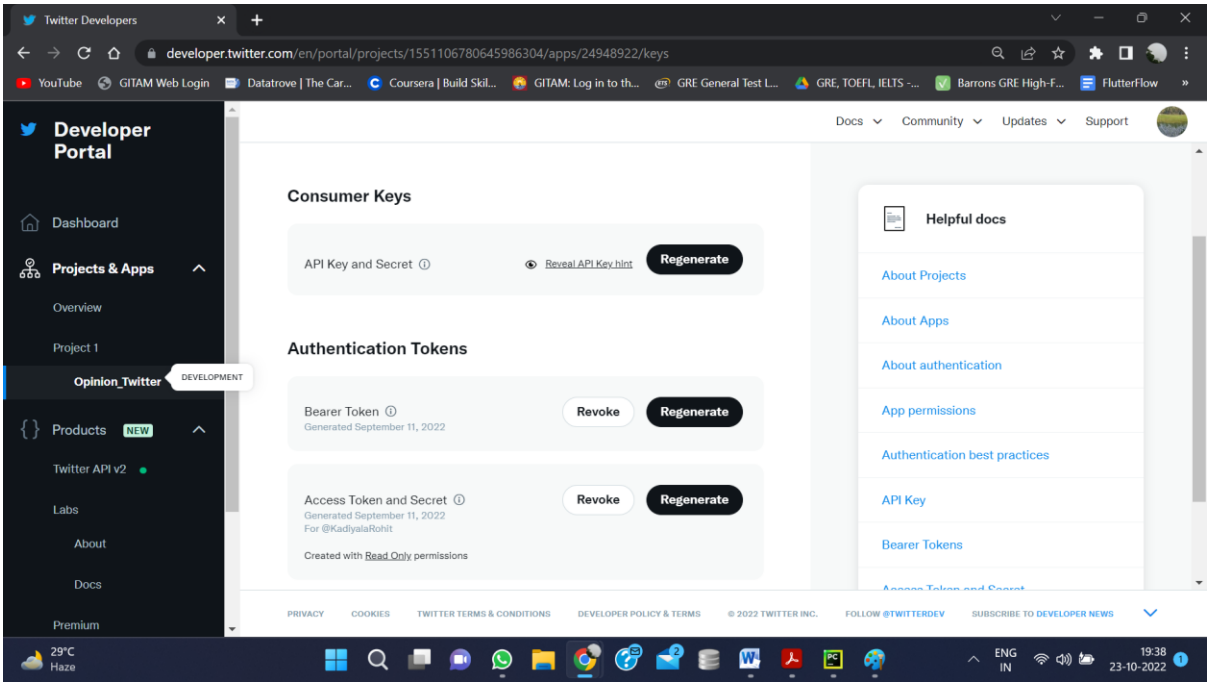
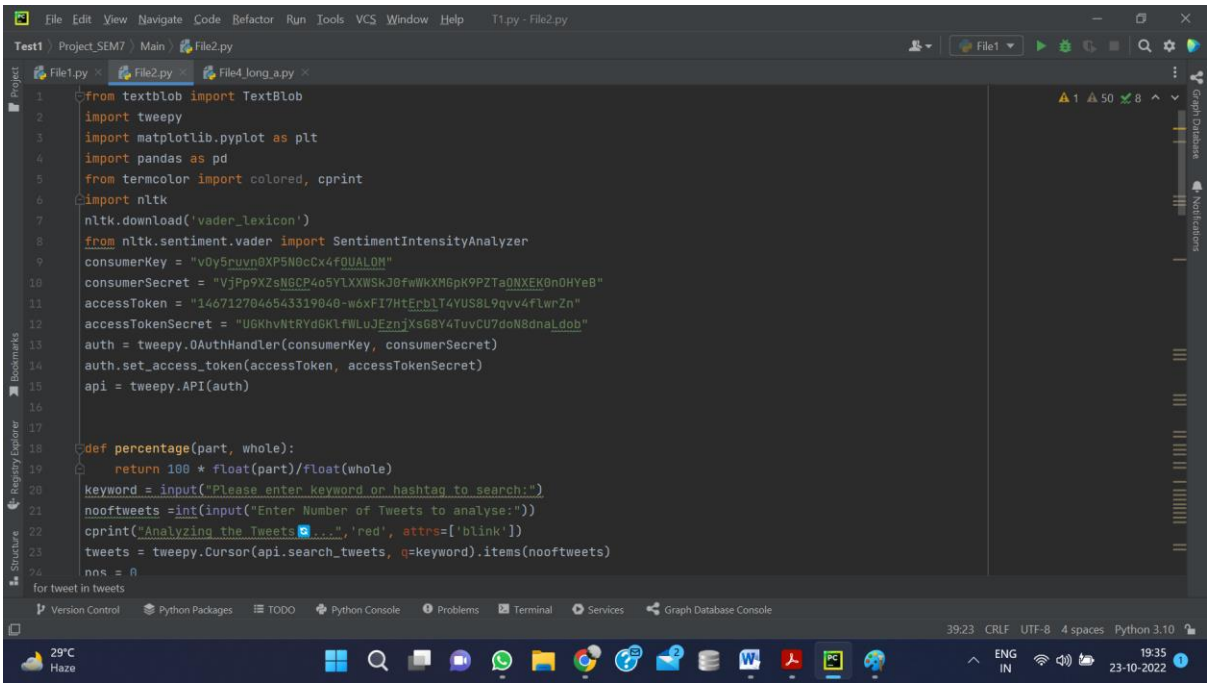
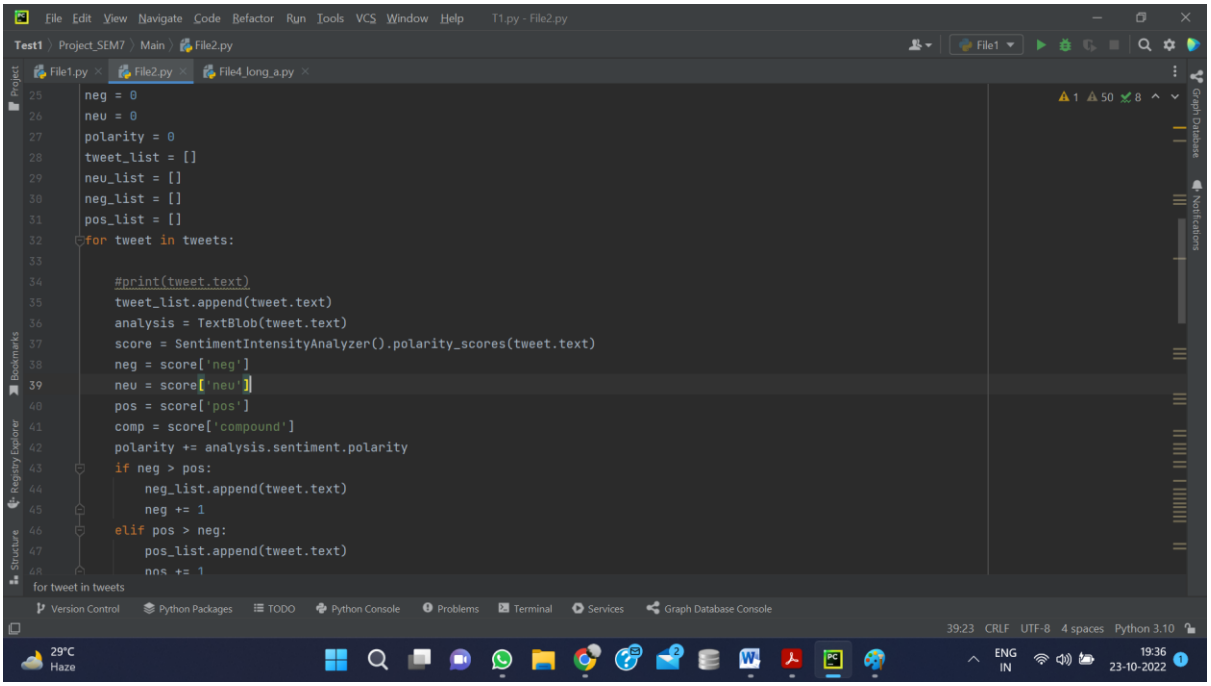


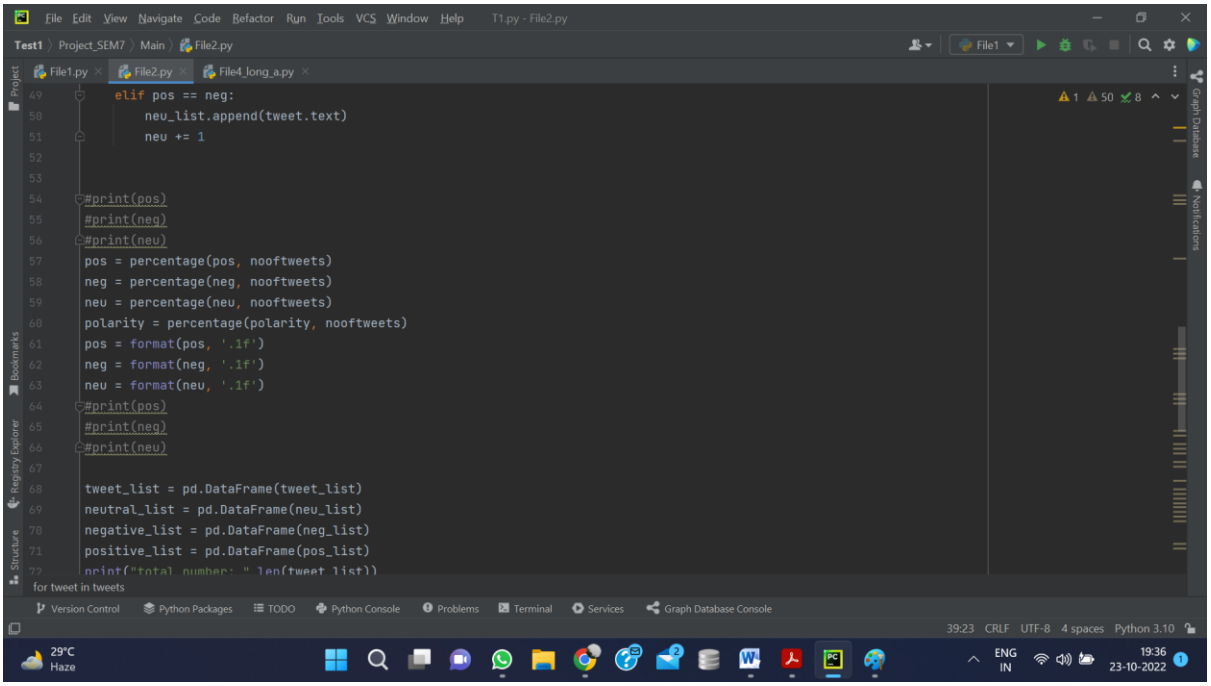
Fig 6.1.1-c Twitter Developer API Keys and Access Tokens

Twitter Sentiment Analysis Python Code in Pycharm IDE

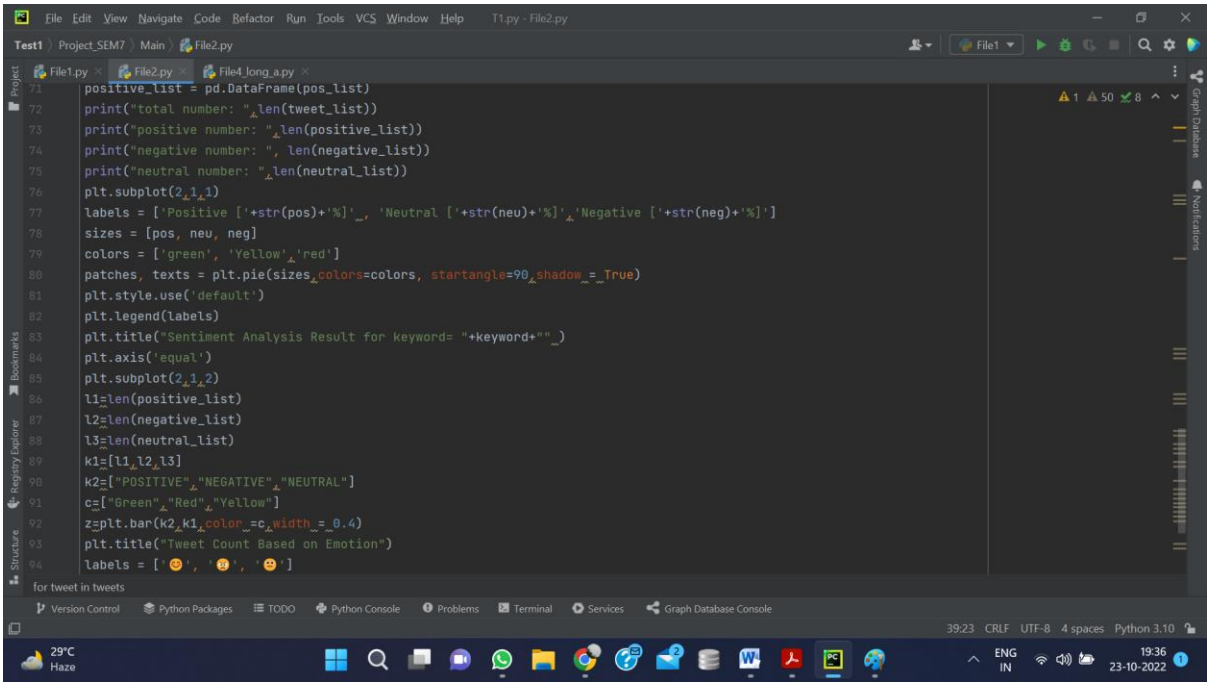


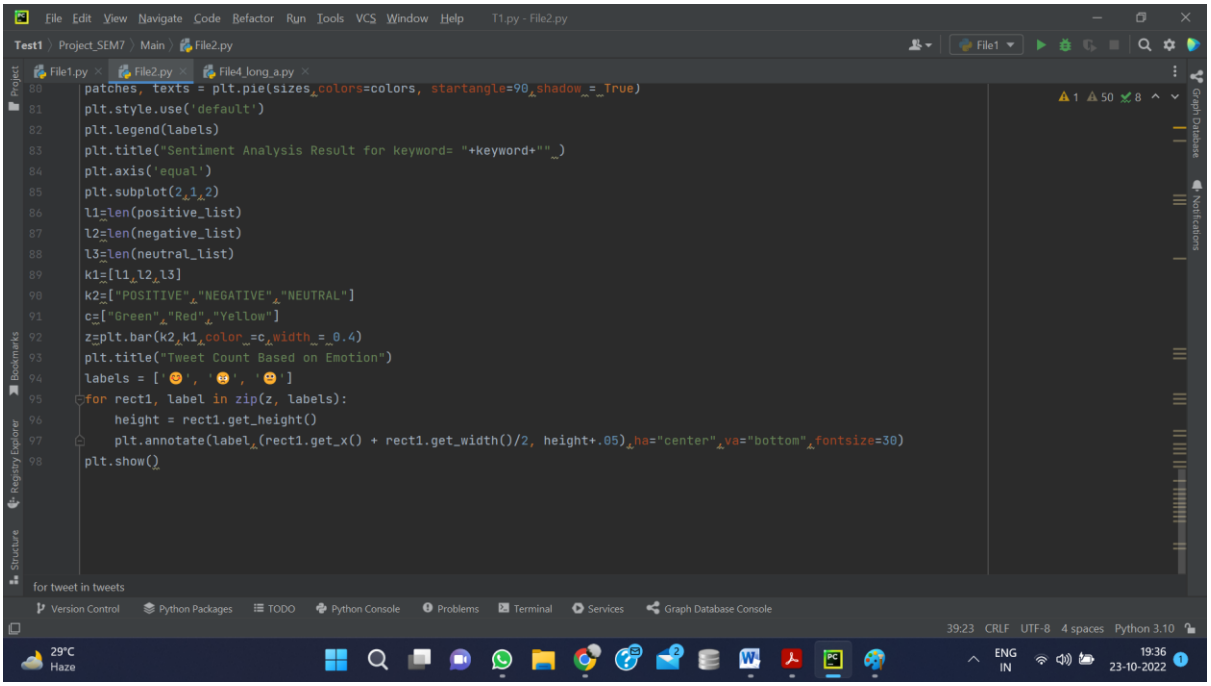


```
neg = 0
neu = 0
polarity = 0
tweet_list = []
neu_list = []
neg_list = []
pos_list = []
for tweet in tweets:

    #print(tweet.text)
    tweet_list.append(tweet.text)
    analysis = TextBlob(tweet.text)
    score = SentimentIntensityAnalyzer().polarity_scores(tweet.text)
    neg = score['neg']
    neu = score['neu']
    pos = score['pos']
    comp = score['compound']
    polarity += analysis.sentiment.polarity
    if neg > pos:
        neg_list.append(tweet.text)
        neg += 1
    elif pos > neg:
        pos_list.append(tweet.text)
        nos += 1
```



```
    elif pos == neg:
        neu_list.append(tweet.text)
        neu += 1


#print(pos)
#print(neg)
#print(neu)
pos = percentage(pos, nooftweets)
neg = percentage(neg, nooftweets)
neu = percentage(neu, nooftweets)
polarity = percentage(polarity, nooftweets)
pos = format(pos, '.1f')
neg = format(neg, '.1f')
neu = format(neu, '.1f')
#print(pos)
#print(neg)
#print(neu)

tweet_list = pd.DataFrame(tweet_list)
neutral_list = pd.DataFrame(neu_list)
negative_list = pd.DataFrame(neg_list)
positive_list = pd.DataFrame(pos_list)
print("total number: " len(tweet_list))
```





6.1.2 For News Article Prediction

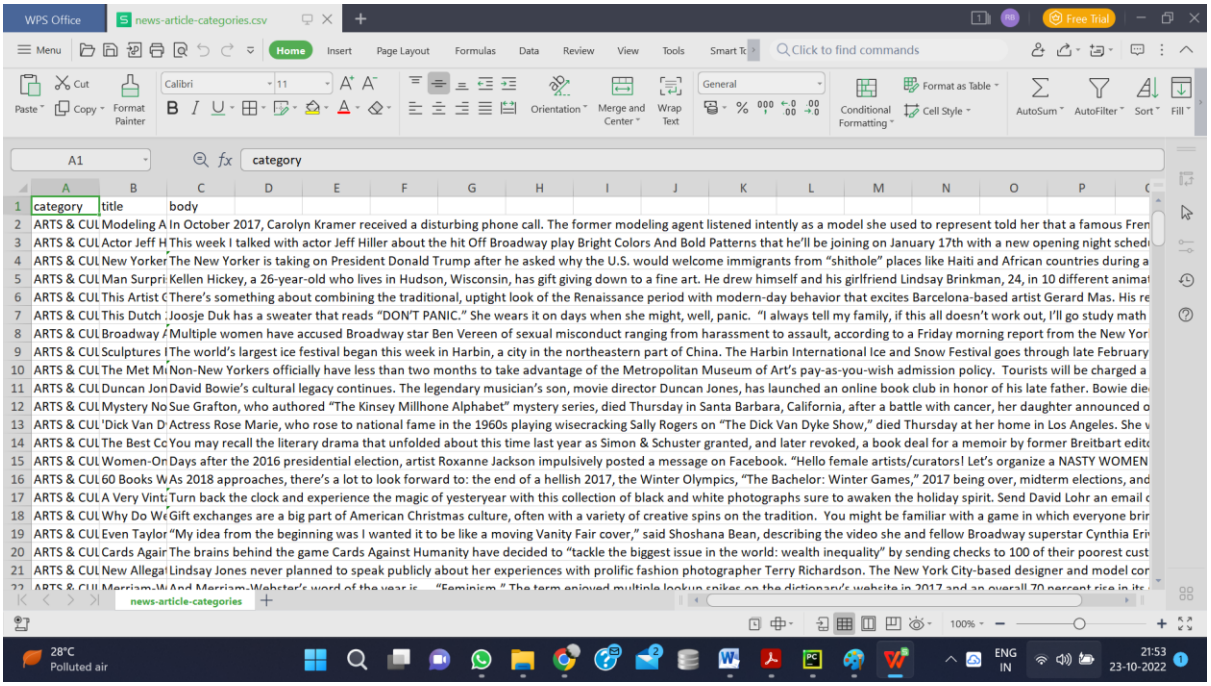
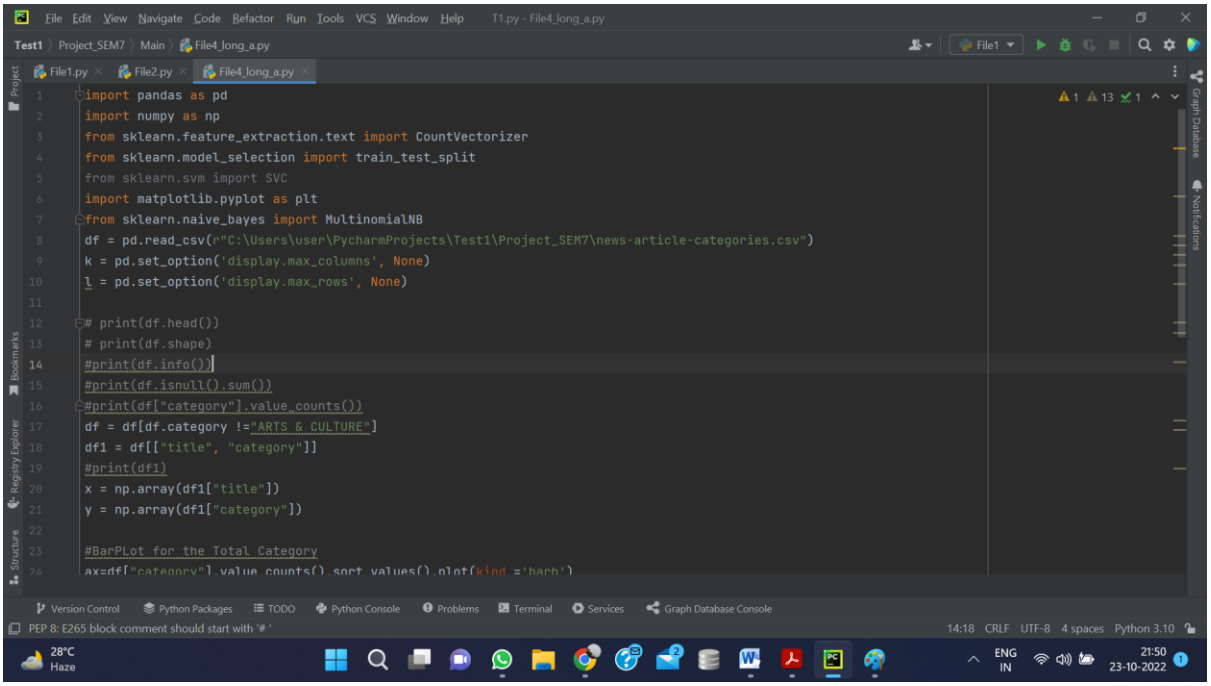
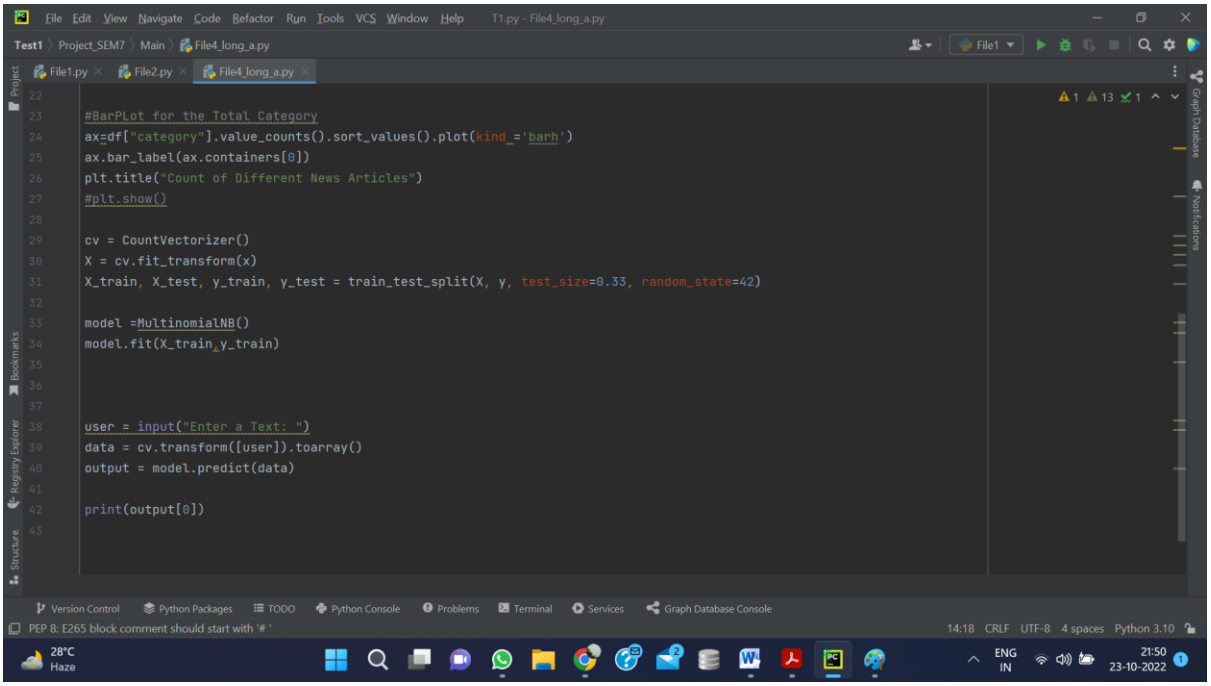
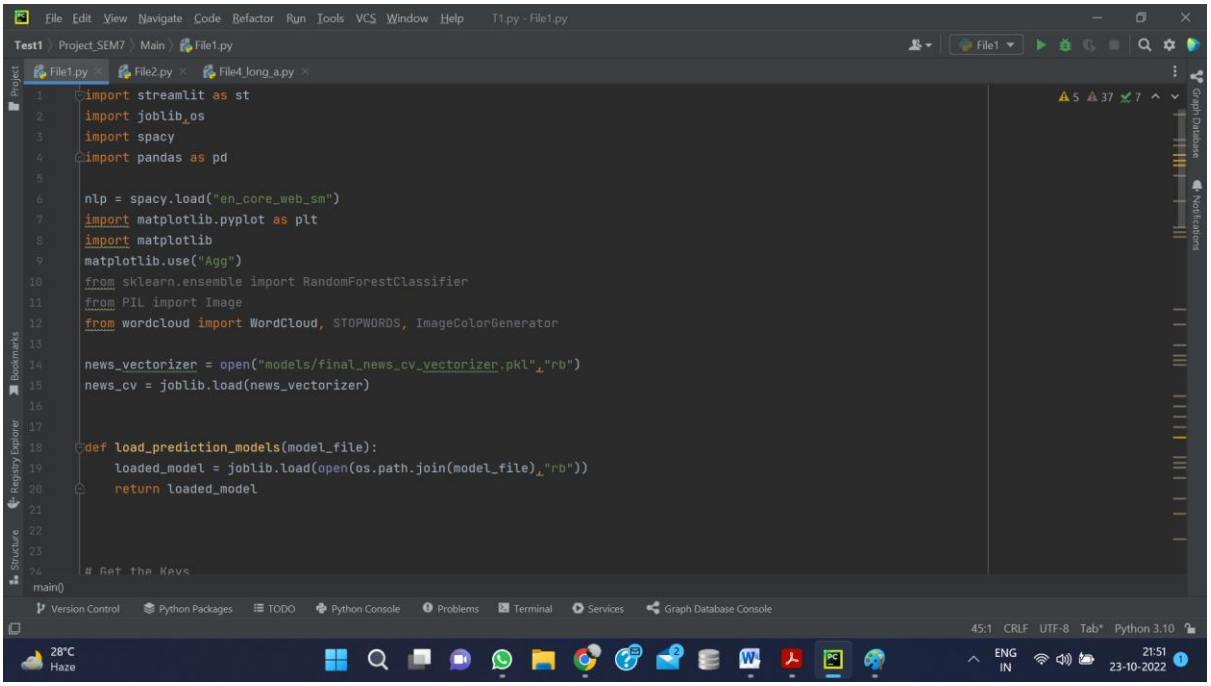


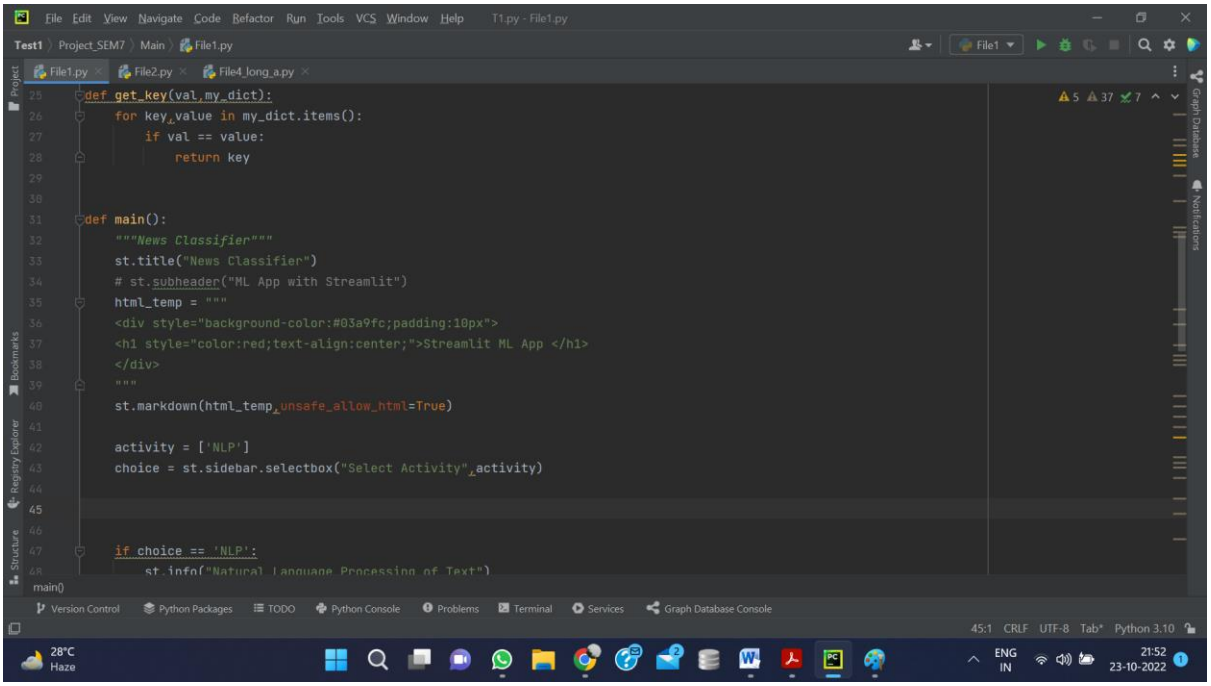
Fig-a: Dataset CSV File



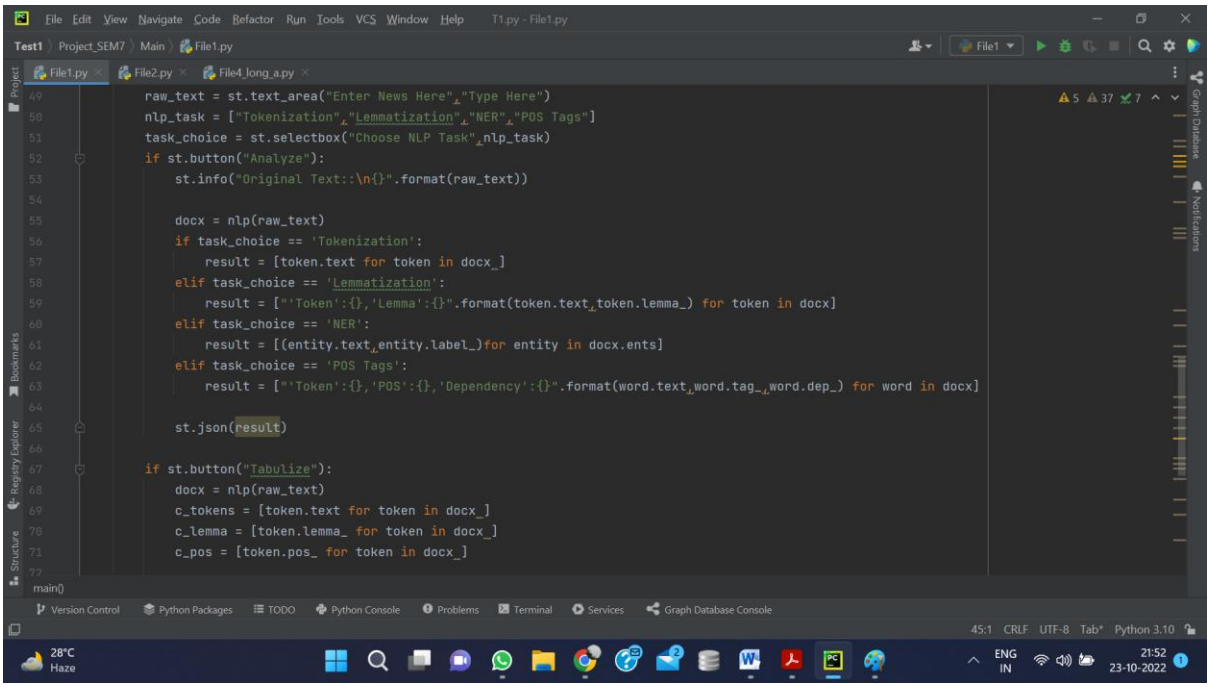


6.1.3 For NLP with Streamlit

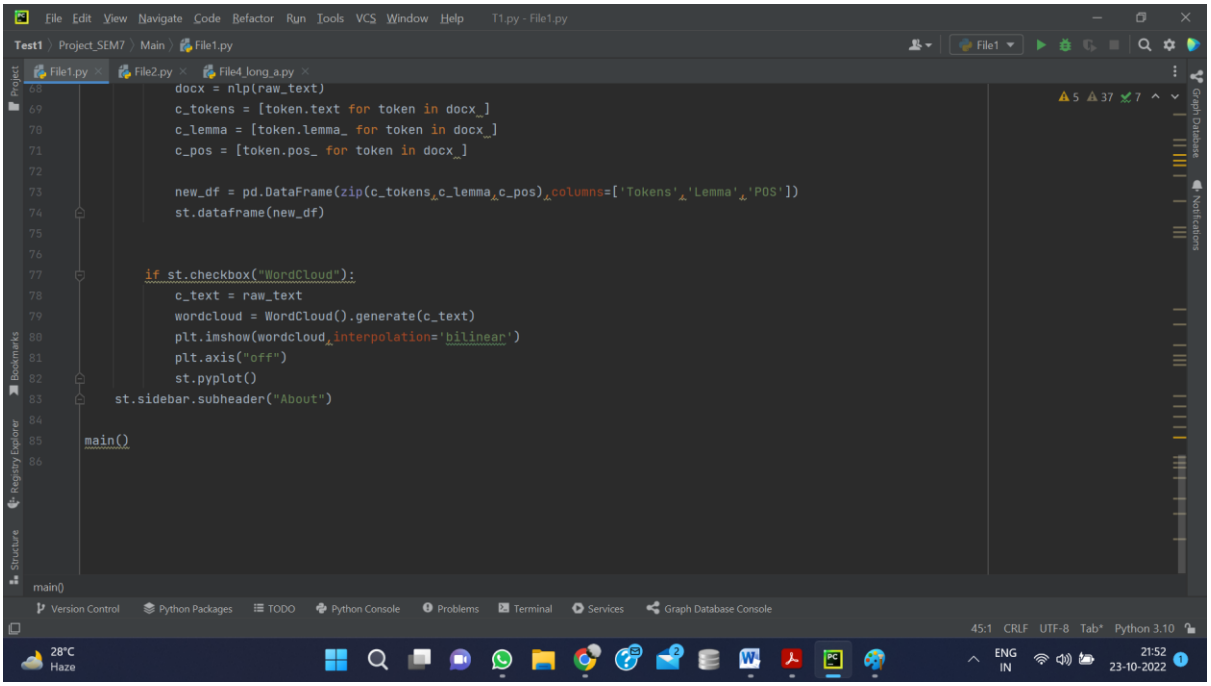




```
def get_key(val,my_dict):
    for key,value in my_dict.items():
        if val == value:
            return key


def main():
    """News Classifier"""
    st.title("News Classifier")
    # st.subheader("ML App with Streamlit")
    html_temp = """
    <div style="background-color:#03a9fc;padding:10px">
    <h1 style="color:red;text-align:center;">Streamlit ML App </h1>
    </div>
    """
    st.markdown(html_temp,unsafe_allow_html=True)

    activity = ['NLP']
    choice = st.sidebar.selectbox("Select Activity",activity)


    if choice == 'NLP':
        st.info("Natural Language Processing of Text")
```



```
        raw_text = st.text_area("Enter News Here","Type Here")
        nlp_task = ["Tokenization","Lemmatization","NER","POS Tags"]
        task_choice = st.selectbox("Choose NLP Task",nlp_task)
        if st.button("Analyze"):
            st.info("Original Text::\n{}".format(raw_text))

            docx = nlp(raw_text)
            if task_choice == 'Tokenization':
                result = [token.text for token in docx]
            elif task_choice == 'Lemmatization':
                result = ["'Token':{},'Lemma':{}".format(token.text,token.lemma_) for token in docx]
            elif task_choice == 'NER':
                result = [(entity.text,entity.label_)for entity in docx.ents]
            elif task_choice == 'POS Tags':
                result = ["'Token':{},'POS':{},'Dependency':{}".format(word.text,word.tag_,word.dep_) for word in docx]

            st.json(result)

        if st.button("Tabulize"):
            docx = nlp(raw_text)
            c_tokens = [token.text for token in docx]
            c_lemma = [token.lemma_ for token in docx]
            c_pos = [token.pos_ for token in docx]
```



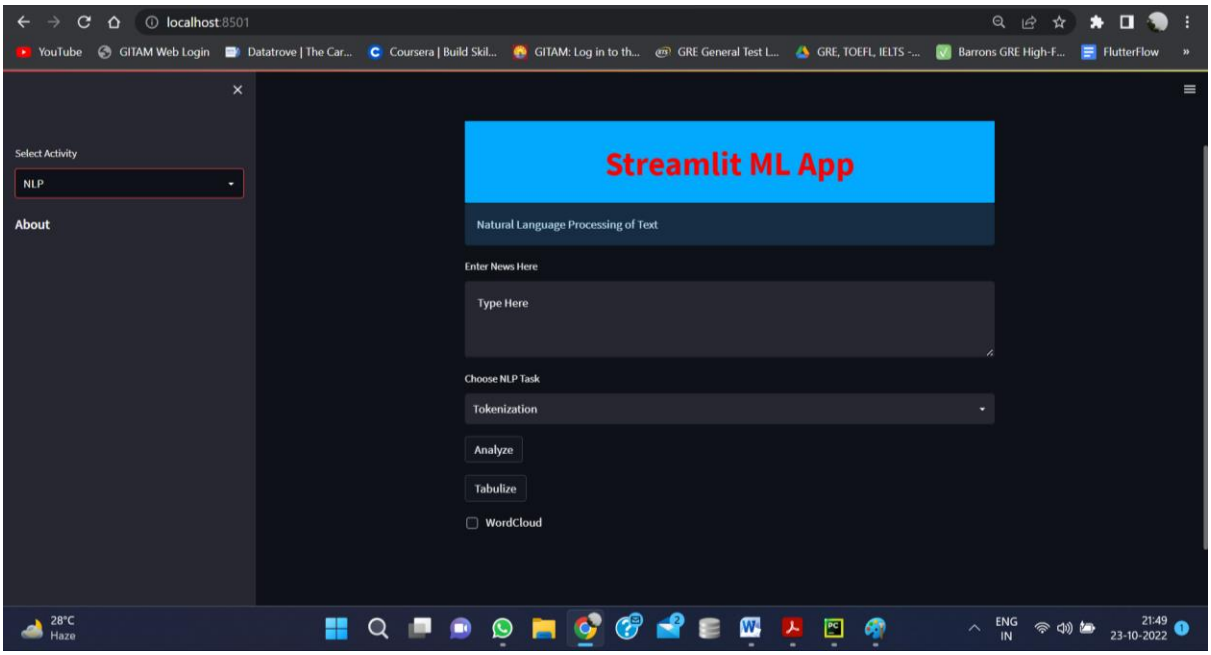
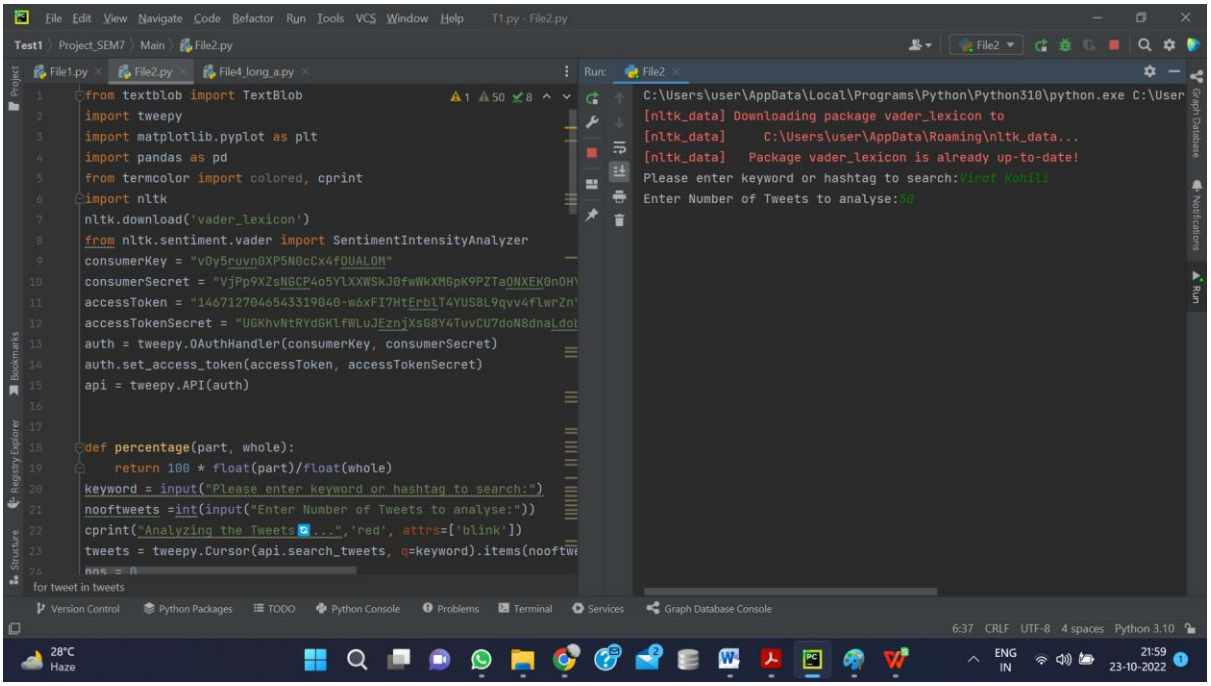
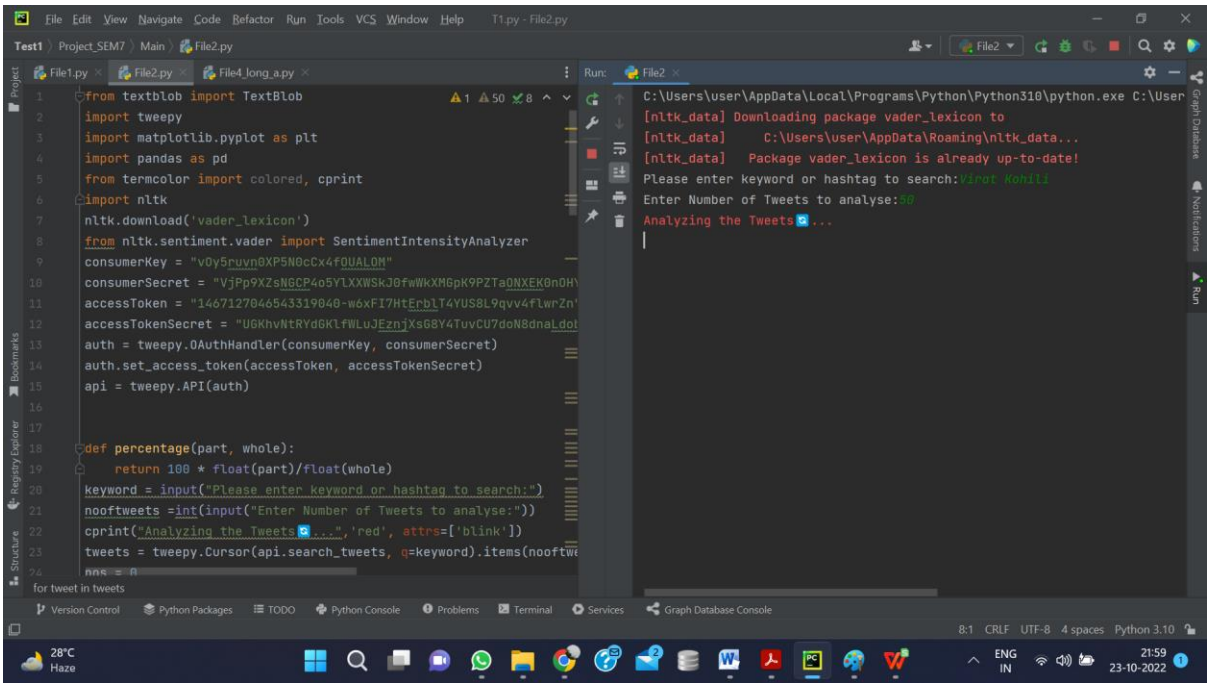


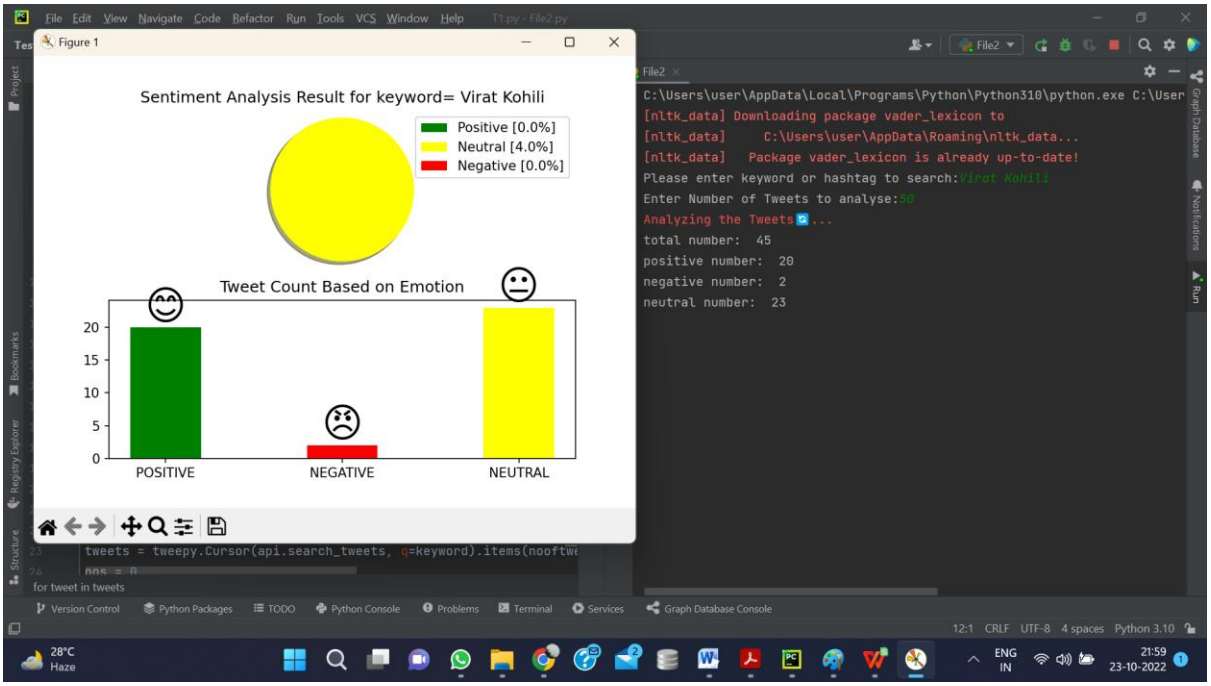
Fig-a: Streamlit Front End Interface

7. RESULTS & DISCUSSION

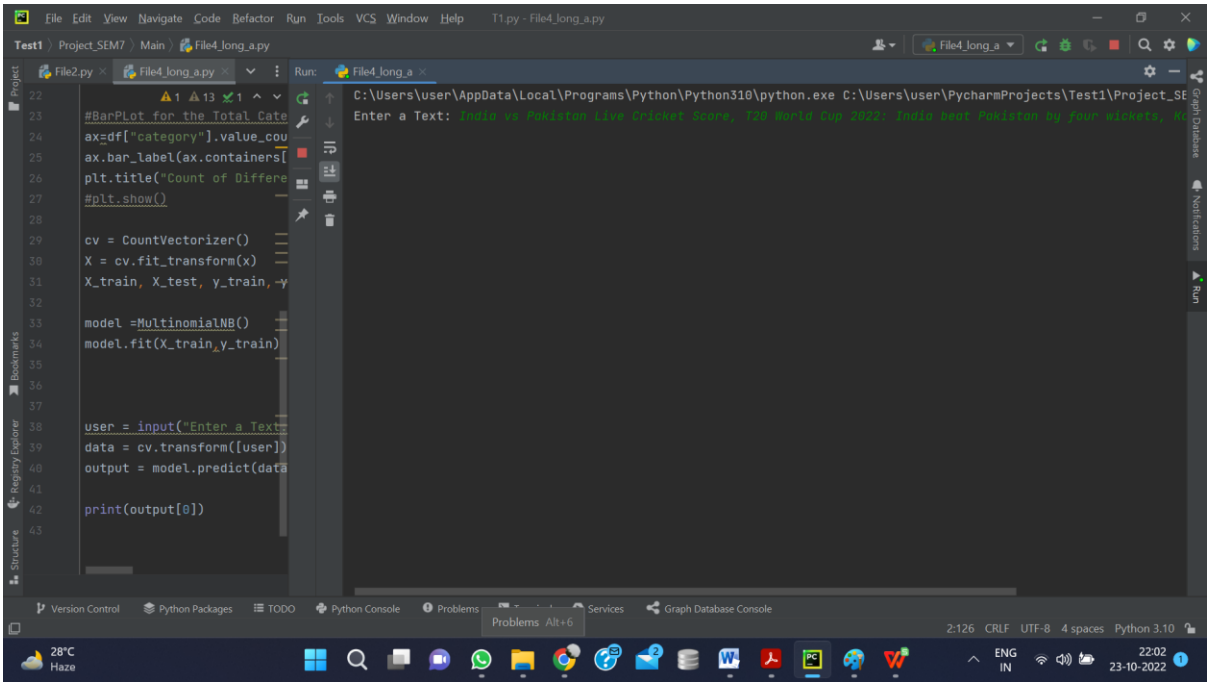
7.1 Twitter Sentiment Analysis Output

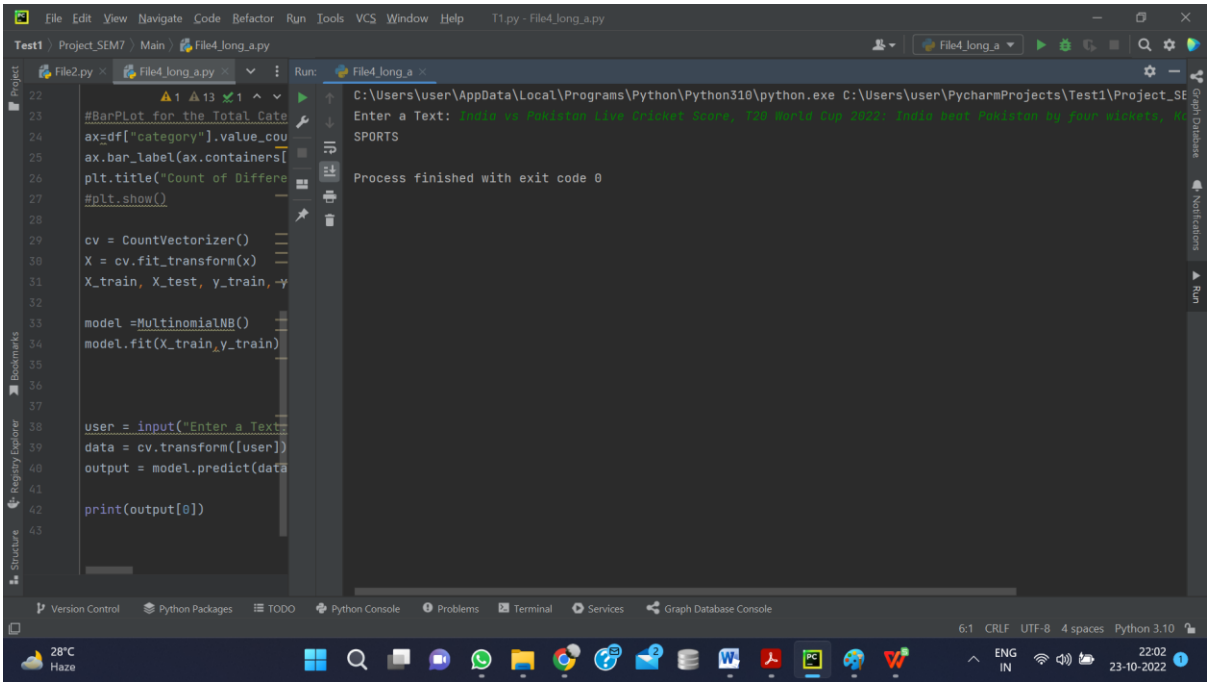




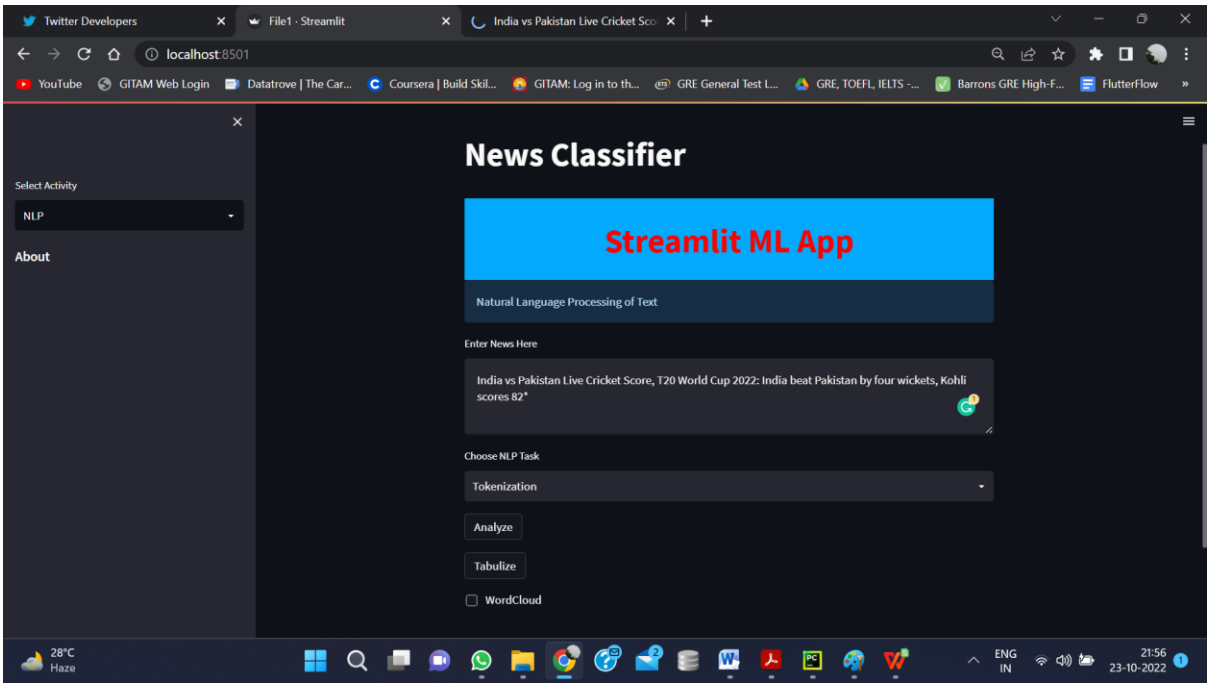


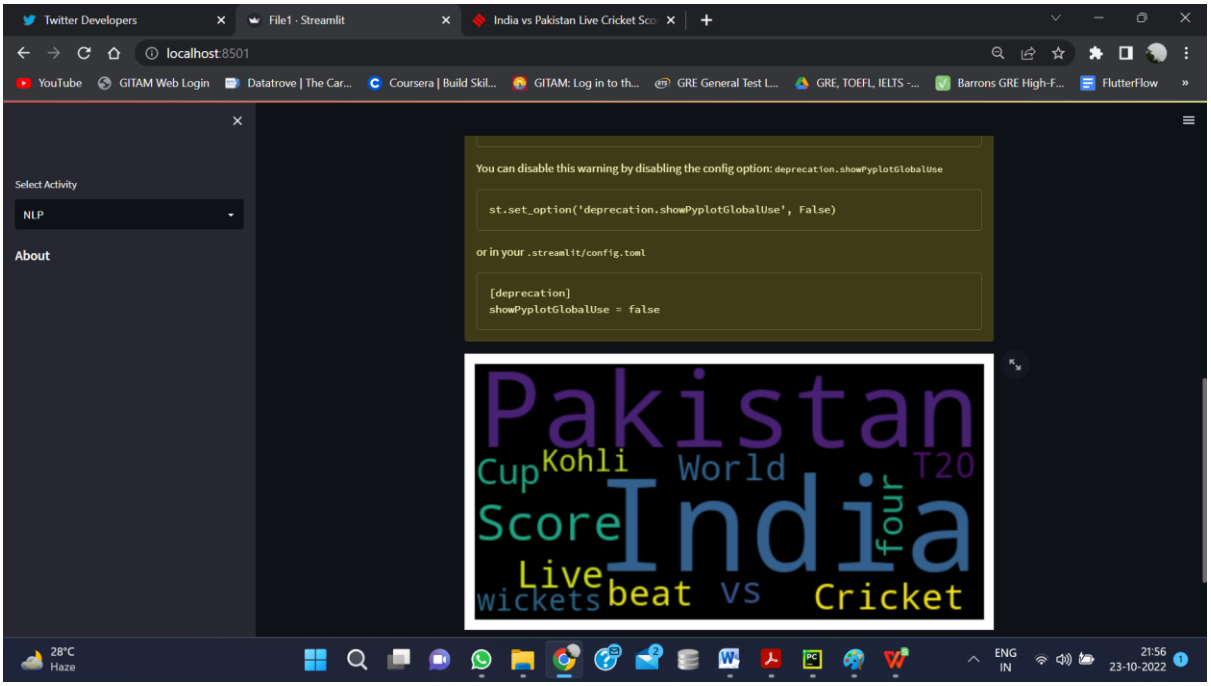
7.2 News Article Prediction with ML Output

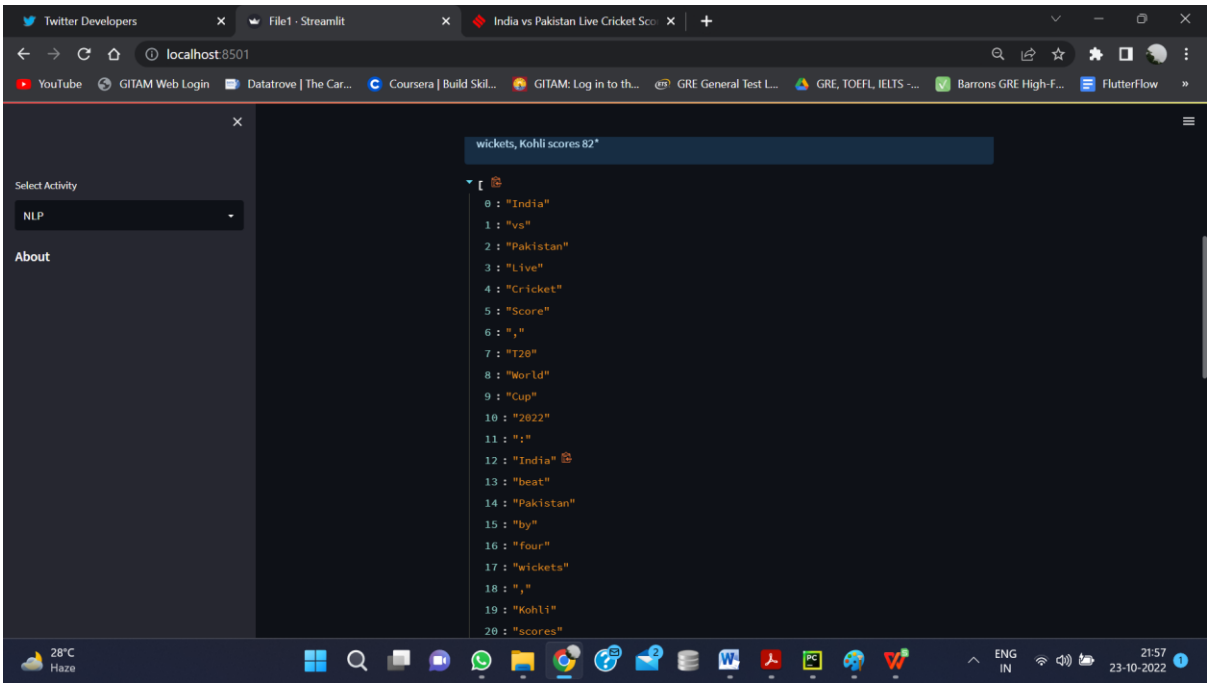


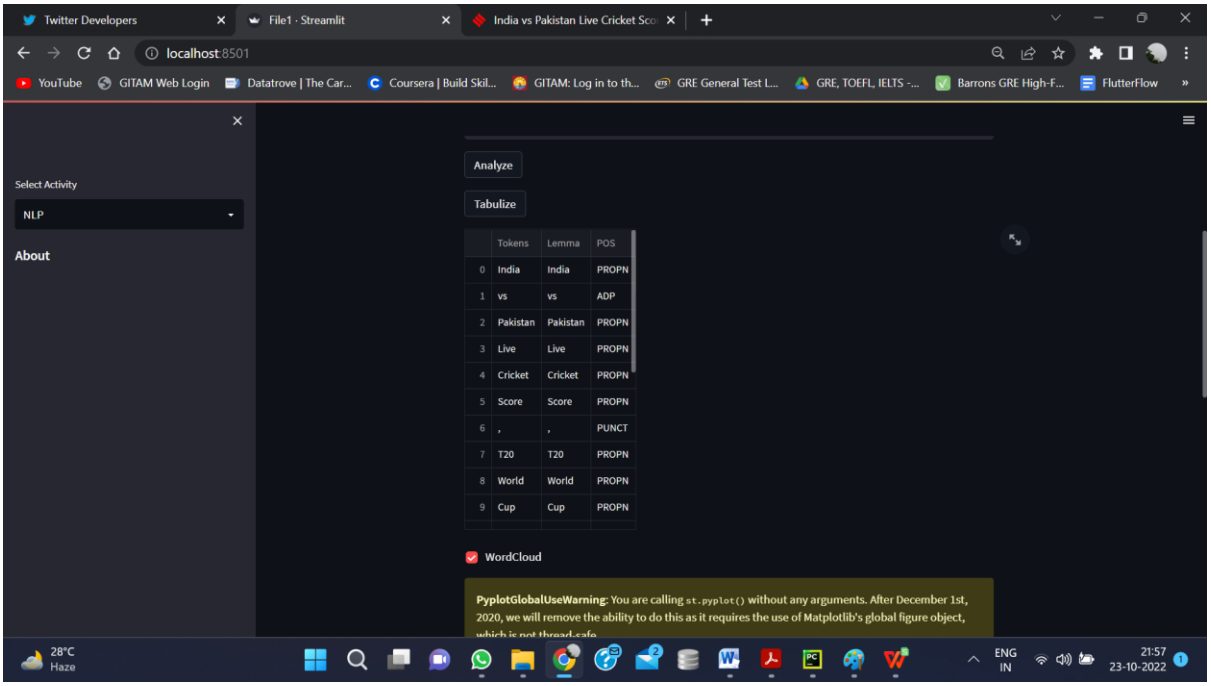


7.3 Streamlit Output:




Twitter Developers
File1 · Streamlit
India vs Pakistan Live Cricket Sco
localhost:8501
YouTube
GITAM Web Login
Datatrove | The Car...
Coursera | Build Skil...
GITAM: Log in to th...
GRE General Test L...
GRE, TOEFL, IELTS -...
Barrons GRE High-F...
FlutterFlow
Select Activity
NLP
About
You can disable this warning by disabling the config option: deprecation.showPyplotGlobalUse
st.set_option('deprecation.showPyplotGlobalUse', False)
or in your .streamlit/config.toml
[deprecation]
showPyplotGlobalUse = false
Pakistan
Cup
Kohli
World
T20
Score
India
four
Live
beat
vs
wickets
Cricket
28°C
Haze
ENG
IN
21:56
23-10-2022


Twitter Developers
File1 · Streamlit
India vs Pakistan Live Cricket Sco
localhost:8501
YouTube
GITAM Web Login
Datatrove | The Car...
Coursera | Build Skil...
GITAM: Log in to th...
GRE General Test L...
GRE, TOEFL, IELTS -...
Barrons GRE High-F...
FlutterFlow
Select Activity
NLP
About
wickets, Kohli scores 82*
0 : "India"
1 : "vs"
2 : "Pakistan"
3 : "Live"
4 : "Cricket"
5 : "Score"
6 : ","
7 : "T20"
8 : "World"
9 : "Cup"
10 : "2022"
11 : ":"
12 : "India"
13 : "beat"
14 : "Pakistan"
15 : "by"
16 : "four"
17 : "wickets"
18 : ","
19 : "Kohli"
20 : "scores"
28°C
Haze
ENG
IN
21:57
23-10-2022


Twitter Developers
File1 · Streamlit
India vs Pakistan Live Cricket Sco
localhost:8501
YouTube
GITAM Web Login
Datatrove | The Car...
Coursera | Build Skil...
GITAM: Log in to th...
GRE General Test L...
GRE, TOEFL, IELTS -...
Barrons GRE High-F...
FlutterFlow
Select Activity
NLP
About
Analyze
Tabulize
Tokens Lemma POS
0 India India PROPN
1 vs vs ADP
2 Pakistan Pakistan PROPN
3 Live Live PROPN
4 Cricket Cricket PROPN
5 Score Score PROPN
6 , , PUNCT
7 T20 T20 PROPN
8 World World PROPN
9 Cup Cup PROPN
WordCloud
PyplotGlobalUseWarning: You are calling st.pyplot() without any arguments. After December 1st, 2020, we will remove the ability to do this as it requires the use of Matplotlib's global figure object, which is not thread-safe
28°C
Haze
ENG
IN
21:57
23-10-2022

8. CONCLUSION & FUTURE SCOPE

Twitter Sentiment Analysis Conclusion

Finally, we exhibited a system for analysing textual Twitter data in this Project, with an emphasis on the growing topic of sentiment analysis. The application will use a more accurate machine-based learning approach for sentiment analysis, as well as natural language processing approaches. As a consequence, the program's emotion will be classified into positive, negative, and neutral categories, which will be displayed in the Matplotlib window by a pie chart and a bar graph. These outcomes matter a lot for sales marketing companies and many MNC's for improving their profits and work force

Future Scope in this Project

A large amount of work needs to be done; here, we present a glimmer of hope for future research directions.

Sarcasm Interpretation: At the moment, the suggested approach is incapable of interpreting sarcasm. The use of irony to mock or show contempt is known as sarcasm. Sarcasm reverses the polarity of an ostensibly positive or negative speech in the context of current work. Conclusion and Future Work "A comprehensive understanding of the basics of discourse-driven sentiment analysis can overcome this constraint." This approach's major purpose is to scientifically discover lexical and pragmatic characteristics that distinguish between ironic, positive, and negative word use.

Multi-lingual support: Developing a multi-language lexical dictionary is presently not viable due to the lack of a multi-lingual lexical dictionary.

- Additional research may be done to make the classifiers language independent.
- The authors developed a sentiment analysis method based on support vector machines; we might adopt a similar technique to make our system language neutral.
- By analysing emoji/smiley feelings to determine neutrality, we may enhance our data gathering and analysis approach.
- Future study can be carried out with potential enhancements such as improved data and a more accurate algorithm.

News Article Prediction and Classification Conclusion

This Project compares the prediction performance of the multi-class category predictor. Using a BBC news dataset with five categories, well-known machine learning methods (Nave Bayes, Random Forest, K-Nearest Neighbour, and Support Vector Machine) were used to build news category predictors (business, sports, technology, politics, and entertainment). The Confusion Matrix was then analysed using performance assessment methods to evaluate the test dataset's Precision, Recall, and Total Accuracy. As a result, the MultinomialNB model was shown to be the most accurate of the four supervised learning models in categorising news articles, with 80% accuracy. The accuracy of the LR model is the lowest (65). However, the performance of the KNN model may be improved by finding the optimal number of neighbours K. The Accuracy of the Model change from dataset to dataset. As the dataset size increases the computational time also increases. This Project also provides and gives the insights of different NLP results for News Text such as Lemmatization, Tokenization, and word cloud also.

This project's working knowledge may be expanded to new use cases, and its accuracy can be improved by utilising numerous prospective technologies in the disciplines of machine learning, data science, and artificial intelligence. We may use the notion of Neural Network to forecast news stories much more correctly. GRU, a Deep Learning idea in development, can also be implemented in the future to eliminate the requirement for time-consuming algos. As the study continues, GRU believes that this technology can also be applied in the sentiment analysis project. In the future, news stories will be automatically categorised and categorised using AI technology that attempts to grasp the structure by analysing its word usage by enlisting the assistance of NLTK

9. REFERENCES

[1] Pang and Lee 2008(Opinion Mining and Sentiment Analysis) Foundations and Trends in Information Retrieval Vol. 2, Nos. 1–2 (2008) 1–135 c2008 B. Pang and L. Lee DOI: 10.1561/1500000001

[2] Hatzivassiloglou and McKeown 1997(Predicting the Semantic Orientation of Adjectives)**.** Proceedings of the 8th Conference on European Chapter of the Association for Computational Linguistics Madrid, Spain, 174-181.");

[3] Esuli and Sebastiani 2006(SENTIWORDNET: A Publicly Available Lexical Resource for Opinion Mining) Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06) Determining Term Subjectivity and Term Orientation for Opinion Mining (11th Conference of the European Chapter of the Association for Computational Linguistics)

[4] U Suleymanov and S Rustamov (Automated News Categorization using Machine Learning methods) IOP Publishing Aegean International Textile and Advanced Engineering Conference (AITAE 2018) IOP Conf. Series: Materials Science and Engineering459 (2019) 012006 doi:10.1088/1757-899X/459/1/012006