International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 04 | Apr 2021 www.irjet.net p-ISSN: 2395-0072

Twitter Sentiment Analysis using Tweepy

Jayesh Chaudhary[1], S. Niveditha[2]

[1]*Computer Science and Engineering, SRM Institute of Science and Technology, Chennai, India*
[2]*Assistant Professor, Computer Science and Engineering, SRM Institute of Science and Technology, Chennai, India*

Abstract - *Sentiment analysis has been present in the civilizations long before the dawn of Technology advancements. But now with the advancements in technology and more & more things being put up on the Internet. The need for Sentiment analysis has increased at a rapid pace. Sentiment analysis has become one of the important tools of the companies and influential figures. Sentiment analysis is being carried out using preset dataset which in turn take a long time to be created and have their own difficulties such as overfitting and bias. Complexity is an another issue in creating a dataset since higher the complexity of problem, larger the amount of data required. In the current fast paced world where things change in an instant, sometimes using a dataset to probe sentiment analysis might result in stale or inaccurate result .That's why we came up with a system which will use live data instead of using a dataset. Sentiment analysis should always be accurate and in correlation with the current views and conditions. We are using Tweepy API as our way to access twitter database and Textblob as our analyzer. Using Tweepy API as our access way will allow us to get live and real time tweets and allow us to perform real time sentiment analysis. By this way we will be able to remove the hassle of creating a dataset and storing it.*

1. OBJECTIVE

To create a system using Tweepy API and Textblob with their necessary terms that will not require a dataset to perform sentiment analysis. With this system we will be able to draw a huge amount of tweets.

2. LITERATURE SURVEY

2.1 Title: Twitter Sentiment analysis using Novelty Detection

Description: This paper uses the approach of which collects large amounts of data and from that data set it excludes the redundant tweets. For this the principle of Novelty detection is used and k-means clustering algorithm and Naive bayes algorithm. K-means algorithm is used to find the outliers in the dataset.[1]

2.2 Title: Sentiment analysis and opinion mining applied to Scientific papers

Description: This study aims to evaluate the people's views, beliefs about a certain topic or others. This uses NLP and machine learning. The purpose of this study is to evaluate review orientation automatically. To solve this problem a hybrid solution is suggested.[2]

3. EXISTING SYSTEM:

Current systems require a preset dataset to be present in order to perform sentiment analysis. Creating a dataset requires a huge amount of resources and time with the need for larger data size increasing with the complexity of the problem. There are various types of machine learning techniques such as Supervised and Unsupervised learning which are used to perform sentiment analysis.

Some of them are mentioned below:

1. Neural Networks
2. Rule Based Classifiers
3. Lexicon and Corpus -based
4. Probablisitic Classifiers
5. Support vector Machines(SVM)

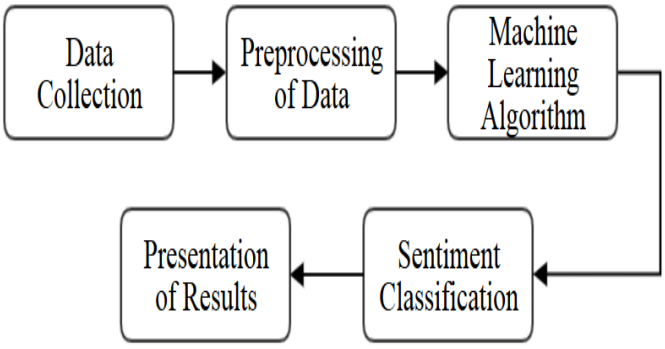

Fig 1 Existing System Architecture

4. PROPOSED SYSTEM

In this paper, we are proposing a system to perform sentiment analysis with a more easier approach which

does not require to create a dataset beforehand. We will be using Twitter Developer with Tweepy API to access the tweets and use Textblob Library.

5. SYSTEM ARCHITECTURE

5.1 Twitter Database

5.2 Tweepy API

5.3 Sentiment Analyzer

5.1. TWITTER DATABASE:

Through twitter database we will be able to access the tweets being made

5.2.TWEEPY API:

Python Library which gives a convenient way to access Twitter API with its own set of classes and methods.

5.3 SENTIMENT ANALYZER:

Performs sentiment analysis on filtered tweets using Naive Bayes Algorithm and Textblob and Regex

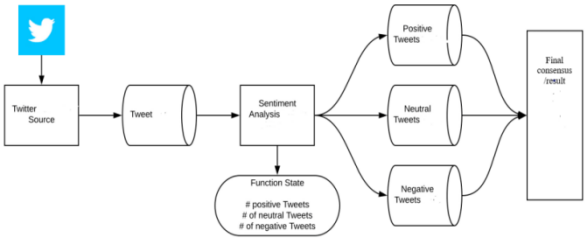

Fig 2. Architecture Diagram

6. PREREQUISITES

- Atom
- Python3
- Visual Studio Code
- Tweepy API
- Twitter Developer Access
- Textblob

6.1.ATOM :

A lightweight, open source and cross platform code editor

6.2. PYTHON3

An interpreted high-level general-purpose programming language. Python's design philosophy emphasizes code readability with its notable use of significant indentation

6.3. VISUAL STUDIO CODE:

A freeware source code editor and compiler with cross platform features for Linux, Windows and MacOS including support for debugging, snippets, code completion.

6.4.TWEEPY API:

An open source python package which makes it convenient to use Twitter API with its own sets of classes and methods.

6.5. TWITTER DEVELOPER ACCESS/PROFILE:

A series of self service software that allows developers to access ,maintain and build their projects and applications.

6.6. TEXTBLOB:

Python library for processing textual data.

7. MODULE

7.1.STREAMING THE TWEETS : Streaming the tweets means streaming and drawing the data live from the database or server. Streaming in one sense means getting the data as simultaneously it is updated. The data streamed is usually up-to-date which allows the system to keep up with the current trend. In this project this is the method which has been implemented.

Import Stream Listener Class and OauthHandler. Create a Class for streaming and processing live tweets.

Define listener which will handle the authentication and connection to Twitter Streaming API. Create another listener that just prints received tweets to stdout.

Authenticate using config.py and connect to Twitter Streaming API. Streamed Tweets are stored in a JSON file

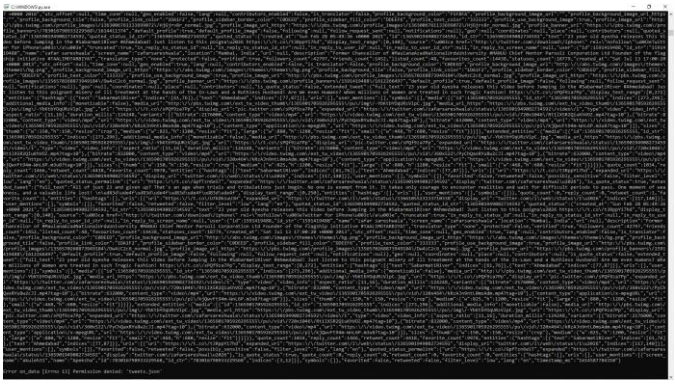

Fig.3: Terminal consisting of Streamed Tweets

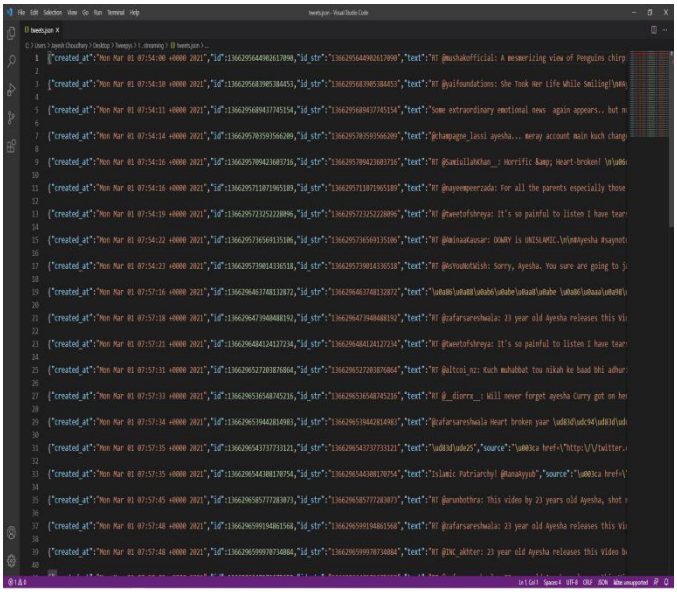

Fig 4: JSON file containing Tweets

7.2 ACCESSING THE TWEETS:

Since the tweets have been streamed, they have to be accessed in a way it does not disrupt the quality of the data and allows us to run the algorithms on it.

The method used in this system is implemented as below: In this module, we will access the tweets we have streamed and stored beforehand. To access the Tweets, we will have to import Cursor and API from TweepyCreate a new Twitter Class for OauthHandler and Stream Listener Class Authenticate using conig.py.

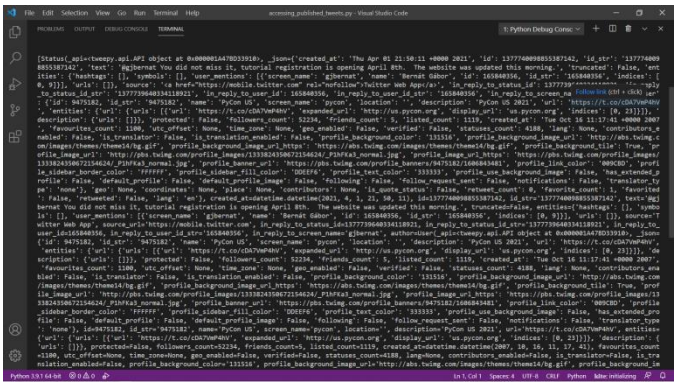

Fig 5. Accessing Tweets in Terminal

7.3. ANALYZING THE TWEETS :

The method used in this project is defined below In this part we will analyse the Twitter Data we have gathered before.

For this, first we will import numpy and pandas. Create a new dataframe for storing the details.

Create a new class Class Analyzer for analysing the tweets. It will separate the tweets on the basis of their different attributes.

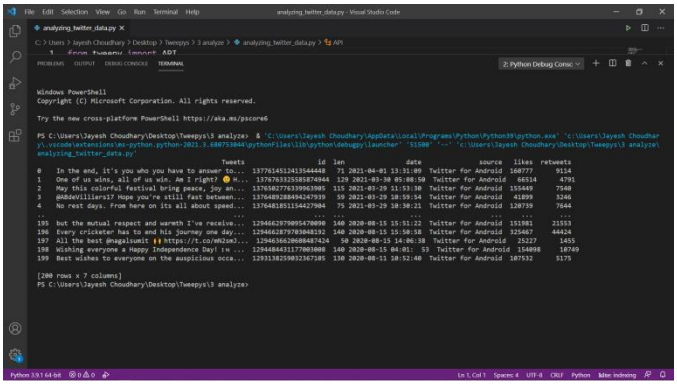

Fig 6 . Analyzing the data

7.4. VISUALIZING THE TWEETS:

Digital representation of facts and statistics is known as data visualisation. Data analysis applications make it easy to see and understand trends, outliers, and regularities by using graphic elements including tables, graphs, and maps.

The method used in this project is explained below: In this part we will visualize the data related to certain tweets and their reaction. For this we will import matplotlib for the graph after analyzing the tweets we will be able to plot graph and layered time series.

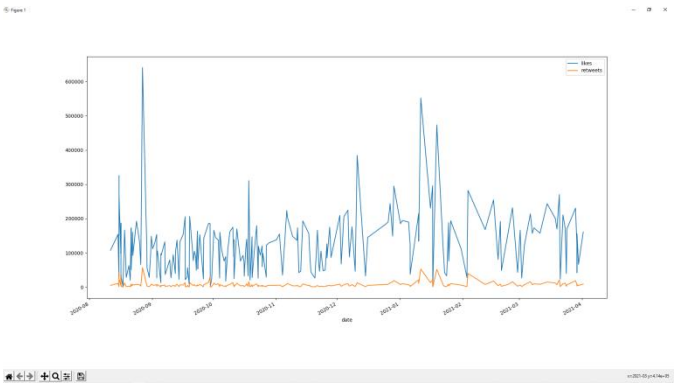

Fig 7. Plot graph B/w RT and time

7.6 SENTIMENT ANALYSIS:

Sentiment analysis is the systematic identification, extraction, quantification, and study of affective states and subjective knowledge using natural language processing, text analysis, computational linguistics, and biometrics.

In this part we will perform sentiment analysis on the Twitter Data we have gathered before.

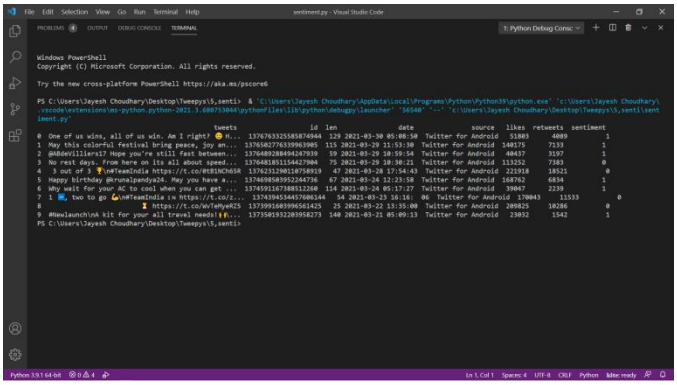

Fig 8 Sentiment analysis with polarity

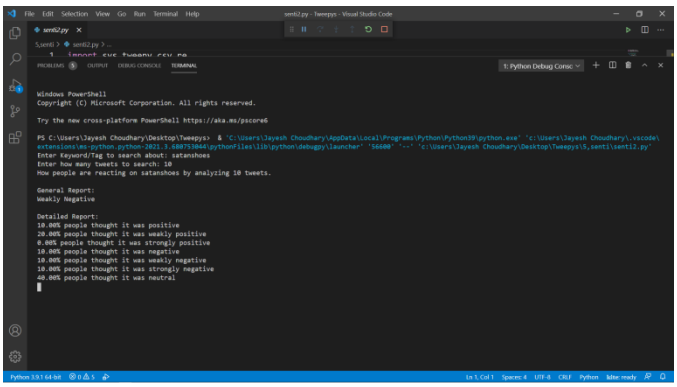

Fig 9 Sentiment analysis with percentage

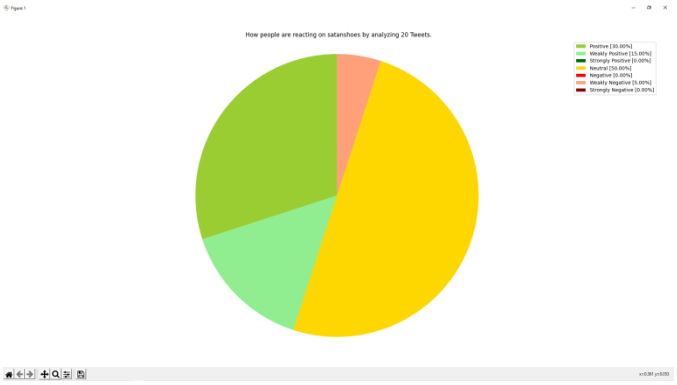

Fig 10 Pie graph showing percentage

8. RESULTS:

We were successfully able to stream, analyze, visualize the data. After completing the aforementioned steps we were able to perform sentiment analysis successfully

9. REFERENCES

1. Twitter Sentiment Analysis using Novelty Detection-Mahalakshmi Shanmugam, Anchal Tiwari -Proceedings of the Third International Conference on Smart Systems and Inventive Technology (ICCSIT 2020)

2. Sentiment analysis and opinion mining applied to scientific paper reviews-Intelligent Data Analysis 23 (2019) 191–214 191 DOI 10.3233/IDA-173807 IOS Press-Brian Keith Norambuena∗, Exequiel Fuentes Lettura and Claudio Meneses Villegas Department of Computing and Systems Engineering, Universidad Católica del Norte, Coquimbo, Chile

3. Sahar A. El_Rahman, " Sentiment Analysis of Twitter Data", Computer and Information sciences College Princess Nourah Bint Abdulrahman University

4. Anurag P. Jain, "Sentiments Analysis Of Twitter Data Using Data Mining", Dept. of Information Technology Pimpri Chinchwad College of Engineering Pune, India,2015 ICIP.

5. Rasika Wagh & Payal Punde , "Survey on Sentiment Analysis using Twitter Dataset",(Mtech): Department of Computer Science and IT, Dr. BAMU Aurangabad, India, ICECA 2018.

6. Adyan Marendra Ramadhani & Hong Soon Goo, "Twitter Sentiment Analysis using Deep Learning Methods", Department of Management Information Systems Dong-A University Busan South Korea, 2017.

7. Bing Liu, Sentiment Analysis and Opinion Mining Morgan and Claypool Publishers, May 2012.

8. V. Kharde and S. Sonawane, "Sentiment Analysis of Twitter Data: A Survey of Techniques", International Journal of Computer Applications, vol. 139, p. 11, 2016.

9. Huma Parveen & Prof. Shikha Pandey "Sentiment Analysis on Twitter Data-set using Naive Bayes Algorithm", Dept. of Computer Science and Engineering Rungta College of Engineering and Technology Bhilai, India , 2016.