



TWITTER SENTIMENT ANALYSIS

Arya Patil^a, Bhakti Koshti^a, Shweta Vibhute^a, Preeti Powar^a, A. R. Ranjane^a.

^aStudent, Sanjay Ghodawat Polytechnic, Atigre, India

Faculty (C.S.E), Sanjay Ghodawat Polytechnic, Atigre, India

ABSTRACT

Sentiment analysis deals with identifying and classifying opinions or sentiments expressed in source text. Social media is generating a vast amount of sentiment rich data in the form of tweets, status updates, blog posts etc. Sentiment analysis of this user generated data is very useful in knowing the opinion of the crowd. Knowledge base approach and Machine learning approach are the two strategies used for analyzing sentiments from the text. By doing sentiment analysis in a specific domain, it is possible to identify the effect of domain information in sentiment classification. We present a new feature vector for classifying the tweets as positive, negative and extract peoples' opinion about products, companies.

Keywords: PYTHON, Twitter, Sentiment analysis

1. INTRODUCTION

Sentiment analysis can be defined as a process that automates mining of attitudes, opinions, views and emotions from text, speech, tweets and database sources through Natural Language Processing (NLP). Sentiment analysis involves classifying opinions in text into categories like "positive" or "negative" or "neutral".

We follow these 3 major steps in our program:

Authorize twitter API client.

Make a GET request to Twitter API to fetch tweets for a particular query.

Parse the tweets. Classify each tweet as positive, negative or neutral.

Hybrid approach. Hybrid sentiment analysis models are the most modern, efficient, and widely-used approach for sentiment analysis. Sentiment analysis is performed by using techniques like Natural Language Processing (NLP), Machine Learning, Text Mining and Information Theory and Coding, Semantic Approach

2. LITERATUREREVIEW

Opinion mining refers to the broad area of natural language processing, text mining, computational linguistics, which involves the computational study of sentiments, opinions and emotions expressed in text. Although, view or attitude based on emotion instead of reason is often colloquially referred to as a sentiment. Hence, lending to an equivalent for opinion mining or sentiment analysis stated that opinion mining has many application domains including accounting, law, research, entertainment, education, technology, politics, and marketing. In earlier days many social media have given web users avenue for opening up to express and share their thoughts and opinions

3. OBJECTIVE AND SCOPE

Twitter sentiment analysis allows you to keep track of what's being said about your product or service on social media, and can help you detect angry customers or negative mentions before they escalate.

Regardless of the name, the goal of sentiment analysis is the same: to know a user or audience opinion on a target object by analyzing a vast amount of text from various sources. You can analyze text on different levels of detail, and the detail level depends on your goals.

The goal of sentiment analysis is the same: to know a user or audience opinion on a target object by analyzing a vast amount of text from various sources. You can analyze text on different levels of detail, and the detail level depends on your goals.

A social media sentiment analysis tells you how people feel about your brand online. Rather than a simple count of mentions or comments, sentiment analysis considers emotions and opinions. It involves collecting and analyzing information in the posts people share about your brand on social media.

4. SCOPE OF PROJECT

Sentiment analysis refers to identifying as well as classifying the sentiments that are expressed in the text source. Tweets are often useful in generating a vast amount of sentiment data upon analysis. These data are useful in understanding the opinion of the people about a variety of topics

The future of sentiment analysis is going to continue to dig deeper, far past the surface of the number of likes, comments and shares, and aim to reach, and truly understand, the significance of social media interactions and what they tell us about the consumers behind the screens.

Tweets are often useful in generating a vast amount of sentiment data upon analysis. These data are useful in understanding the opinion of the people about a variety of topics. on the orders placed and the orders completed.

5. FEATURES OF IDLE

- Cross-platform: works mostly the same on Windows, Unix, and macOS
- Python shell window (interactive interpreter) with colorizing of code input, output, and error messages
- multi-window text editor with multiple undo, Python colorizing, smart indent, call tips, auto completion, and other features
- search within any window, replace within editor windows, and search through multiple files (grep)
- debugger with persistent breakpoints, stepping, and viewing of global and local namespaces
- configuration, browsers, and other dialogs}
- IDLE is Python's Integrated Development and Learning Environment. IDLE has the following features: coded in 100
- Using IDLE is not a requirement for using Python. There are many other IDEs that can be used to write Python programs, not to mention a variety of text-based programmer's editors that many programmers prefer to IDEs.
- IDLE (short for Integrated Development and Learning Environment) is an integrated development environment for Python, which has been bundled with the default implementation of the language since 1.5. 2b1. It is packaged as an optional part of the Python packaging with many Linux distributions.

Usage:

IDLE is used to create, modify, and execute Python code. IDLE provides a fully-featured text editor to write Python scripts and provides features like syntax highlighting, auto-completion, and smart indent. IDLE also has a debugger with features such as stepping and breakpoints.

It allows programmers to easily write Python code. Just like Python Shell, IDLE can be used to execute a single statement and create, modify, and execute Python scripts. IDLE provides a fully-featured text editor to create Python scripts that include features like syntax highlighting, auto completion, and smart indent.

Python IDLE gives you the ability to create and edit these files with ease. Python IDLE also provides several useful features that you'll see in professional IDEs, like basic syntax highlighting, code completion, and auto-indentation, (left or center).

6. DEVELOPMENT

6.1 Python

Python is a computer programming language often used to build websites and software, automate tasks, and conduct data analysis. Python is a general-purpose language, meaning it can be used to create a variety of different programs and isn't specialized for any specific problems.

Python is a general-purpose coding language—which means that, unlike HTML, CSS, and JavaScript, it can be used for other types of programming and software development besides web development. Python can be used for things like: Back end (or server-side) web and mobile app development.

PyCharm, a proprietary and Open Source IDE for Python development. PyScripter, Free and open-source software Python IDE for Microsoft Windows. PythonAnywhere, an online IDE and Web hosting service. Python Tools for Visual Studio, Free and open-source plug-in for Visual Studio.

Python is an interpreter, interactive, object-oriented programming language. It incorporates modules, exceptions, dynamic typing, very high level dynamic data types, and classes. It supports multiple programming paradigms beyond object-oriented programming, such as procedural and functional programming.

Users can quickly make video games from scratch using Python because it is an easy programming language to learn. Python is used to provide add-on game functionality, create 3D graphics, and script various parts of games.

7. UML DIAGRAMS

7.1 Flowchart:

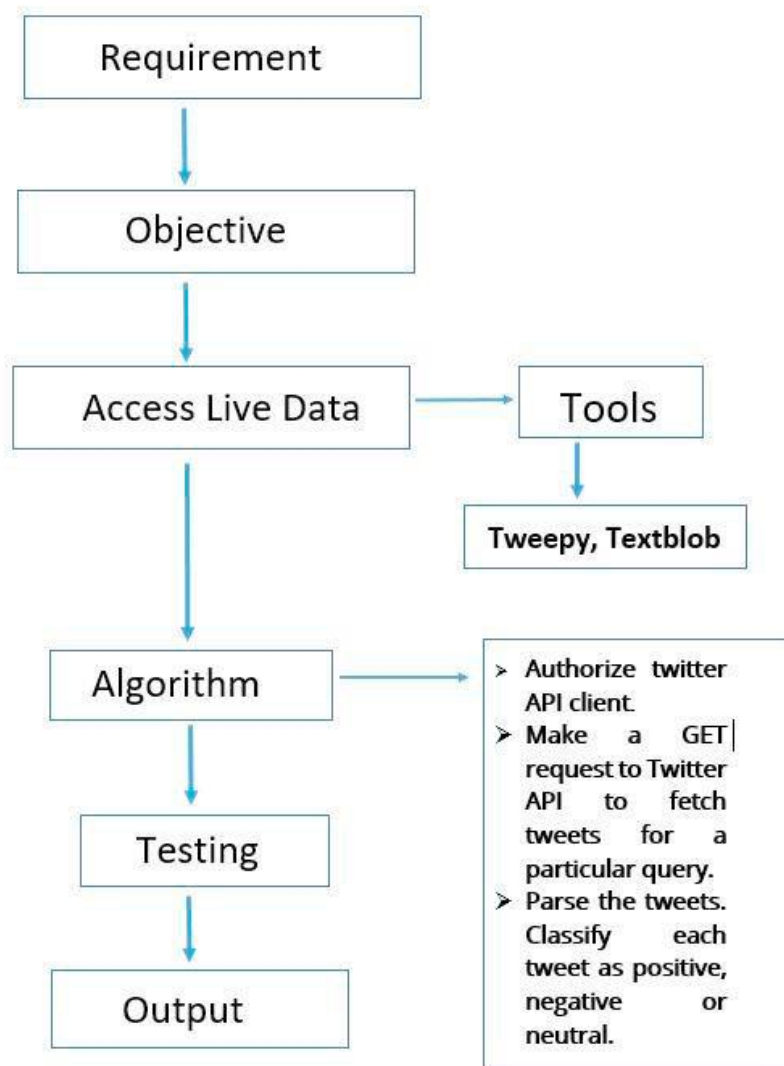
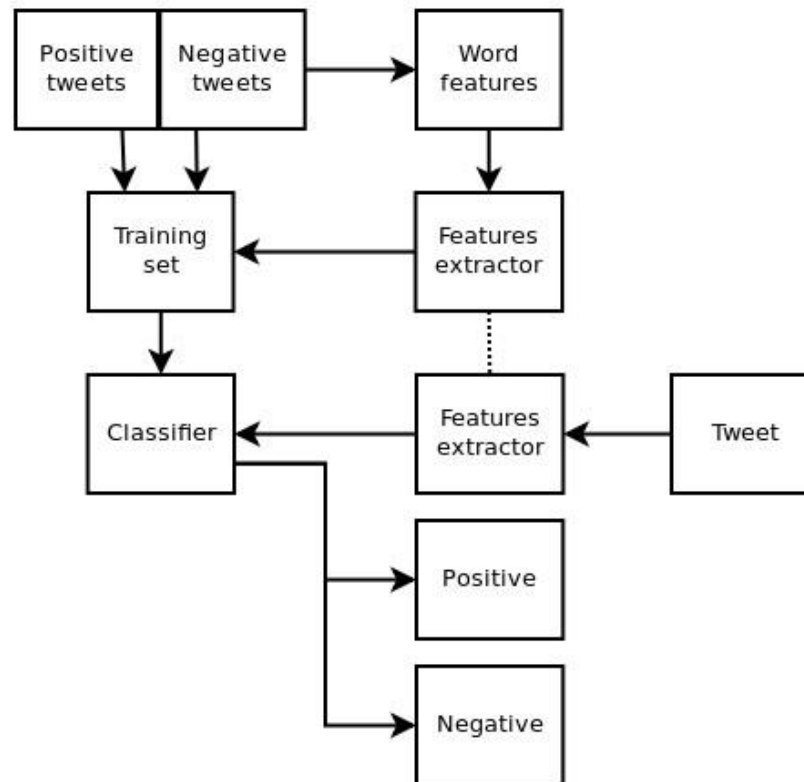


Figure 7.1 Flowchart

7.2 Data Flow Diagrams:**7.2 Data Flow Diagram****8. SYSTEM CODING**

```
import sys,tweepy, csv, re
```

```
from textblob import TextBlob
```

```
import matplotlib.pyplot as plt
```

```
class SentimentAnalysis:
```

```
    def __init__(self):
```

```
        self.tweets = []
```

```
        self.tweetText = []
```

```
    def DownloadData(self):
```

```
        # authenticating
```

```
        consumerKey = '980Whjb0OAY8WwCg7rXKzFbtE'
```

```
        consumerSecret = 'TV0d4fsEGibuhOdVRaBc7WulVxHvZE5E2YptXYwyqyyaTGjXkh'
```

```
accessToken = '744437793305333760-aFYURyw96rdEZqOZj2WSewiuxalwkQ!'
accessTokenSecret = '7YA3qYkIwudAX33P6df55weKmmEINxzm4UN3yzLF6saKG'

    auth = tweepy.OAuthHandler(consumerKey, consumerSecret)
auth.set_access_token(accessToken, accessTokenSecret)

api = tweepy.API(auth)

    # input for term to be searched and how many tweets to search
searchTerm = input("Enter Keyword/Tag to search about: ")
NoOfTerms = int(input("Enter how many tweets to search: "))

    # searching for tweets
self.tweets = tweepy.Cursor(api.search, q=searchTerm, lang = "en").items(NoOfTerms)

    # Open/create a file to append data to
csvFile = open('result.csv', 'a')

    # Use csv writer
csvWriter = csv.writer(csvFile)

    # creating some variables to store info
    polarity = 0
    positive = 0
wpositive = 0
spositive = 0
    negative = 0
wnegative = 0
snegative = 0
    neutral = 0

    # iterating through tweets fetched
for tweet in self.tweets:

    #Append to temp so that we can store in csv later. We use encode UTF-8
self.tweetText.append(self.cleanTweet(tweet.text).encode('utf-8'))

    #tweet's text
```

```
analysis = TextBlob(tweet.text)

# print(analysis.sentiment) # print tweet's polarity

polarity += analysis.sentiment.polarity # adding up polarities to find the average later

if (analysis.sentiment.polarity == 0): # adding reaction of how people are reacting to find average later
    neutral += 1

elif (analysis.sentiment.polarity > 0 and analysis.sentiment.polarity <= 0.3):
wpositive += 1

elif (analysis.sentiment.polarity > 0.3 and analysis.sentiment.polarity <= 0.6):
    positive += 1

elif (analysis.sentiment.polarity > 0.6 and analysis.sentiment.polarity <= 1):
spositive += 1

elif (analysis.sentiment.polarity > -0.3 and analysis.sentiment.polarity <= 0):
wnegative += 1

elif (analysis.sentiment.polarity > -0.6 and analysis.sentiment.polarity <= -0.3):
    negative += 1

elif (analysis.sentiment.polarity > -1 and analysis.sentiment.polarity <= -0.6):
snegative += 1

# Write to csv and close csv file
csvWriter.writerow(self.tweetText)
csvFile.close()

# finding average of how people are reacting
positive = self.percentage(positive, NoOfTerms)
wpositive = self.percentage(wpositive, NoOfTerms)
spositive = self.percentage(spositive, NoOfTerms)
negative = self.percentage(negative, NoOfTerms)
wnegative = self.percentage(wnegative, NoOfTerms)
snegative = self.percentage(snegative, NoOfTerms)
neutral = self.percentage(neutral, NoOfTerms)

# finding average reaction
polarity = polarity / NoOfTerms
```

```
# printing out data
print("How people are reacting on " + searchTerm + " by analyzing " + str(NoOfTerms) + " tweets.")
print()
print("General Report: ")

if (polarity == 0):
    print("Neutral")
elif (polarity > 0 and polarity <= 0.3):
    print("Weakly Positive")
elif (polarity > 0.3 and polarity <= 0.6):
    print("Positive")
elif (polarity > 0.6 and polarity <= 1):
    print("Strongly Positive")
elif (polarity > -0.3 and polarity <= 0):
    print("Weakly Negative")
elif (polarity > -0.6 and polarity <= -0.3):
    print("Negative")
elif (polarity > -1 and polarity <= -0.6):
    print("Strongly Negative")

print()
print("Detailed Report: ")

print(str(positive) + "% people thought it was positive")
print(str(wpositive) + "% people thought it was weakly positive")
print(str(spositive) + "% people thought it was strongly positive")
print(str(negative) + "% people thought it was negative")
print(str(wnegative) + "% people thought it was weakly negative")
print(str(snegative) + "% people thought it was strongly negative")
print(str(neutral) + "% people thought it was neutral")

self.plotPieChart(positive, wpositive, spositive, negative, wnegative, snegative, neutral, searchTerm, NoOfTerms)

def cleanTweet(self, tweet):
```

```
# Remove Links, Special Characters etc from tweet

return ' '.join(re.sub("(@[A-Za-z0-9]+)|(^0-9A-Za-z \t) | (\w +:\/\|\/\S +)", " ", tweet).split())

# function to calculate percentage

def percentage(self, part, whole):

    temp = 100 * float(part) / float(whole)

    return format(temp, '.2f')

def plotPieChart(self, positive, wpositive, spositive, negative, wnegative, snegative, neutral, searchTerm, noOfSearchTerms):

    labels = ['Positive [' + str(positive) + '%]', 'Weakly Positive [' + str(wpositive) + '%]', 'Strongly Positive [' + str(spositive) + '%]', 'Neutral [' + str(neutral) + '%]',

             'Negative [' + str(negative) + '%]', 'Weakly Negative [' + str(wnegative) + '%]', 'Strongly Negative [' + str(snegative) + '%]']

    sizes = [positive, wpositive, spositive, neutral, negative, wnegative, snegative]

    colors = ['yellowgreen', 'lightgreen', 'darkgreen', 'gold', 'red', 'lightsalmon', 'darkred']

    patches, texts = plt.pie(sizes, colors=colors, startangle=90)

plt.legend(patches, labels, loc="best")

plt.title('How people are reacting on ' + searchTerm + ' by analyzing ' + str(noOfSearchTerms) + ' Tweets.')

plt.axis('equal')

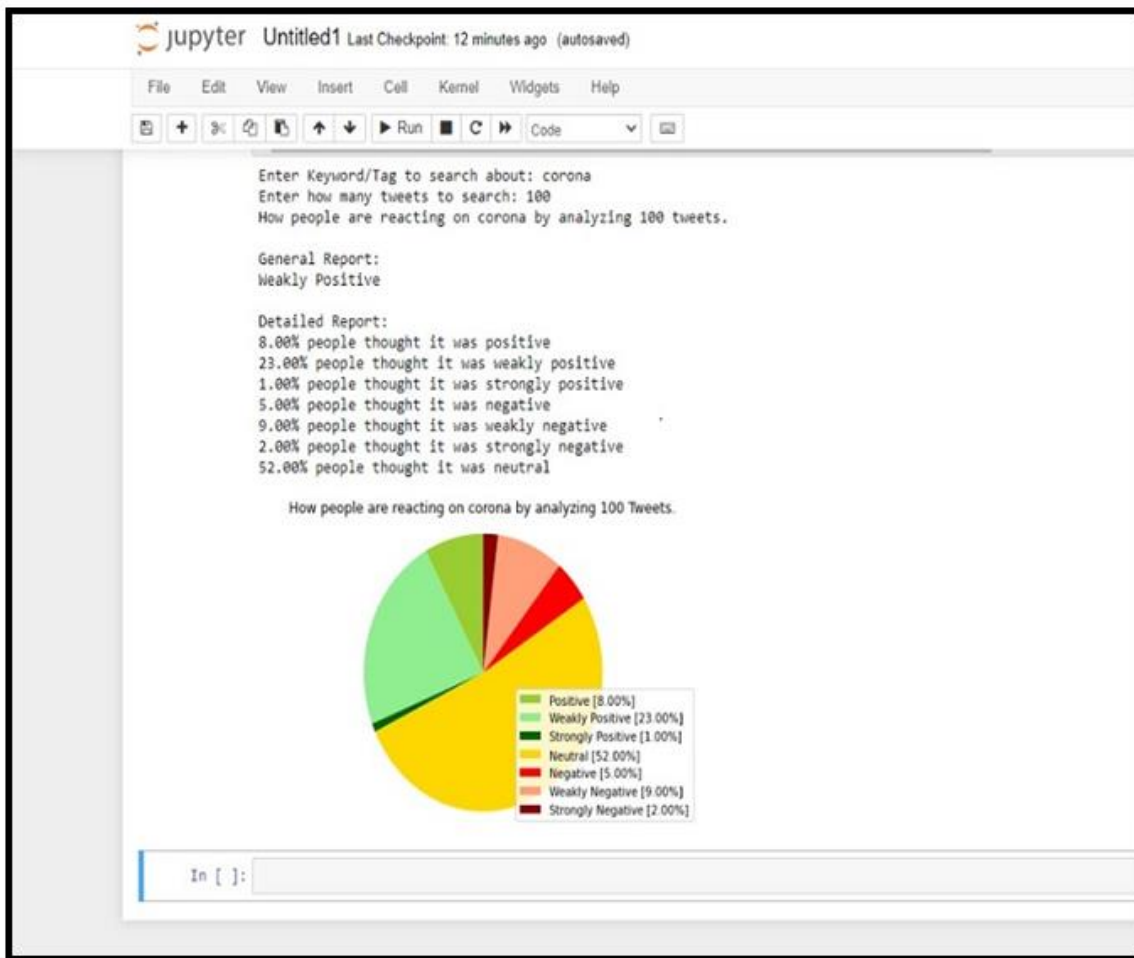
plt.tight_layout()

plt.show()

if __name__ == "__main__":

    sa = SentimentAnalysis()

    sa.DownloadDatabase
```


Screen Shot:**9. SOFTWARE REQUIREMENT**

- Python
- Tweepy: This library allows Python to access the Twitter platform/database using its API.
- TextBlob: It is a Python library for processing textual data. It helps in diving into common natural language processing (NLP) tasks such as part-of-speech tagging, noun phrase extraction, sentiment analysis, classification, translation, and more.
- Matplotlib: Python module used to draw pie chart.

10. ADVANTAGES AND DISADVANTAGES OF PROPOSEDSYSTEM**10.1 Advantages**

1. The sentiment analysis of Twitter data has many organizational benefits such as understanding your brand more deeply, growing your influence, and improving customer service.
2. Sentiment analysis also means you'll be able to detect changes in the overall opinion towards your brand.
3. Sentiment analysis is used to determine whether a given text contains negative, positive, or neutral emotions.
4. Twitter sentiment analysis allows you to keep track of what's being said about your product or service on social media, and can help you detect angry customers or negative mentions before they escalate.

5. Sentiment analysis tools are essential to detect and understand customer feelings. Companies that use these tools to understand how customers feel can use it to improve CX. Sentiment analysis tools generate insights into how companies can enhance the customer experience and improve customer service.

10.2 Disadvantages

1. Limitations of Sentiment Analysis for Reputation management
2. Sentiment Neutral Statements with Negative or Positive Reputational Impact. ...
3. Incorrectly Targeted Sentiment. ...
4. Review Language is Dissimilar to Social Media and News Language.
5. In this article, we talked about popular problems of sentiment analysis classification: sarcasm, negations, word ambiguity, and multipolarity.
6. Limitations of automated sentiment analysis

But computer programs have problems recognizing things like sarcasm and irony, negations, jokes, and exaggerations - the sorts of things a person would have little trouble identifying. And failing to recognize these can skew the results.

11. CONCLUSION

This project is a small step towards the efficient automation of sentiment analysis by focusing on ambiguous statements. The system proposed by us also attempts to extract actual emotions from tweets. Such a system will be very useful for various marketing teams to gain actual and detailed feedback from their users. Also, the social media influencers use sentiment analysis to know about the trends and create the content accordingly. It can be used in production company where they can analyse the success of their movies. In politics, the political parties propose their agenda and can analyse the tweets to decide their strategy for campaigning. In stock market tweets play an important role and hence analysing the sentiments, one can decide whether to invest in a particular share or not.

REFERENCE

- [1] M.Rambocas, and J. Gama, "MarketingResearch:TheRoleofSentimentAnalysis". The 5th SNA-KDD Workshop'11. U niversityofPorto, 201
- [2] M.Rambocas, and J. Gama, "MarketingResearch:TheRoleofSentimentAnalysis". The 5th SNA-KDD Workshop'11. U niversityofPorto, 201
- [3] Y. Zhou, and Y. Fan, " A Sociolinguistic Study of American Slang," Theory and Practice in Language Studies, 3(12), 2209–2213, 2013. doi:10.4304/tpls.3.12.2209-221
- [4] B. Pang, and L. Lee, "Opinion mining and sentiment analysis," 2nd workshop on making sense of Microposts. Ithaca: Cornell University. Vol.2(1), 2008.
- [5] H. Saif, Y. He and H. Alani, "Alleviating Data Scarcity for Twitter Sentiment Analysis". Association for Computational Linguistics, 2012