Contents: Page Numbers

List of Tables Page Number

List of Figures Page Number

1. Introduction

1.1 Briefing

Social Media is an indispensable part of our lives. Mood analysis is basically analyzing the series of statements from a person's post and then depicting the "Mood" of the person.

There are a number of articles that are presented every year in the field of Data Mining and Sentiment Analysis. The number of articles is ever increasing, and this fact creates a need to have surveys to summarize the recent trends in the field for better understanding of the subject.

Cluster analysis is a multivariate method which aims to classify a sample of subjects (or objects) on the basis of a set of measured variables into a number of different groups such that similar subjects are placed in the same group. An example where this might be used is in marketing. It may be useful to identify distinct groups of potential customers so that, for example, advertising can be appropriately targeted.

For the project and to assert to the above fact, we have undertaken a series of experiments and re-implementing the contributions of those experiments. For the simplicity and also the size of the dataset present, we have selected **Twitter** as our social media platform.

1.2. Problem Domain

Data Mining is the current topic in social media and Data Mining through Sentiment Analysis is the domain we are focusing. In the last decade there has been an explosion of interest in the field of data mining. The process of collecting, searching through and analyzing a large amount of data in a database, so that we can discover patterns or relationships is the use of data mining. An upcoming branch in computer science, data mining can be applied to various problems, one of them being sentiment analysis.

Sentiment Analysis or Opinion Mining is the computational study of people's opinion/emotions towards an entity/certain facts and figures.

E.g.: Product reviews in different online shopping sites gives us a better understanding of the product we want to buy.

Sentiment adds to human communications, whether personally or in social media. When captured computationally, customer sentiment expresses beyond facts, that convey emotions and moods.

1.3 Related Studies

There are a number of articles that are presented every year in the field of Data Mining and Sentiment Analysis. The number of articles is ever increasing, and this fact creates a need to have surveys to summarize the recent trends in the field for better understanding of the subject.
For the project and to assert to the above fact, we have undertaken a series of experiments and re-implementing the contributions of those experiments. Although it is not possible to summarize the entire subject, so the basic view and conclusion of the articles have been summarized below.

- Twitter Sentiment Analysis [1]
- Facebook Status Sentiment Analysis [2]
- Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis [3]
- Automatically Predicting mood from expressed emotions [4]

➢ **Twitter Sentiment Analysis**

OBJECTIVE:

The goal for this experiment is two-fold. First, we want to evaluate whether our training data with labels derived from hashtags and emoticons is useful for training sentiment classifiers for Twitter. Second, we want to evaluate the effectiveness of the features from section for sentiment analysis in Twitter data.

CONCLUSION:

Our experiments on twitter sentiment analysis show that part-of- speech features may not be useful for sentiment analysis in the micro blogging domain. More research is needed to determine whether the POS features are just of poor quality due to the results of the tagger or whether POS features are just less useful for sentiment analysis in this domain. Features from an existing sentiment lexicon were somewhat useful in conjunction with micro blogging features, but the Micro blogging features (i.e., the presence of intensifiers and positive/negative/neutral emoticons and abbreviations) were clearly the most useful.

- **Facebook Status Sentiment Analysis**

OBJECTIVE:

Facebook messages are more clearly and briefly explained than other social network sites and are easier to classify than tweets as there is no character or word size.
The objective is to classify the emotions on Facebook status message more accurately by improved information mining via NLP analysis.

CONCLUSION:

It is a very basic research as only binary classification is considered. And the best binary scale would solve harder problems. As many of the expressions have been generalized, chances of misclassification are increased. On the other hand, some basic features of sentiment analysis and text mining is implemented. In a less, hard scale problem this process would be very effective.

- **Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis**

OBJECTIVE:

This paper presents a new approach to phrase-level sentiment analysis that first determines whether an expression is neutral or polar and then disambiguates the polarity of the polar expressions. With this approach, the system is able to automatically identify the contextual polarity for a large subset of sentiment expressions, achieving results that are significantly better than baseline.

CONCLUSION:

In this paper, a new approach to phrase-level sentiment analysis that first determines whether an expression is neutral or polar and then disambiguates the polarity of the polar expressions. With this approach, we are able to automatically identify the contextual polarity for a large subset of sentiment expressions, achieving results that are significantly better than baseline.

➢ **Automatically Predicting mood from expressed emotions**

OBJECTIVE:

In this paper we depict the mood (positive, negative and neutral) of a person by reading the facial expression of person from a video using sensors.

CONCLUSION:

The contributions that this thesis delivered to the field of automatic mood prediction from visual input can be summarized as follows:

- We proposed a novel framework for predicting the mood of a person from video. Our framework is based on the key idea that mood is different from emotion but related to it. Hence, although automatic emotion recognition systems cannot be considered able to predict mood, their output can inform a mood prediction algorithm. Specifically, the emotions expressed by a person over time can help us estimate his/her mood. The advantage of this notion is that it allows reusing the bulk of work on automatic emotion recognition.
- We crafted the first affective database, named EMMA, for the study of automatic mood recognition from video, as well as for the mapping of emotions to mood. The creation of a new database was dictated by the lack of existing audio-visual affective databases targeting mood. Instead, EMMA is appropriate for studying mood prediction in that its videos:
    - last a few minutes each, which is sufficient time for depicting mood and the evolution of more than one emotion,
    - present mood expressed through mild facial and bodily expressions and are annotated in terms of emotion (per frame) and mood (per video).

  The database is publicly available, and it can be a valuable tool for Affective Computing researchers investigating automatic emotion and/or mood recognition from facial and bodily expression.

1.4 Glossary

*Table 1:**Glossary***

Term	Definition
Data Mining	It is the computing process of discovering patterns in large data sets.
Dataset	Most commonly a data set corresponds to the contents of a single database table, in this case a table with keywords under every superset of mood.
Sentiment Analysis	It refers to the use of natural language processing, text analysis, computational linguistics, and biometrics to systematically identify, extract, quantify, and study affective states and subjective information.
TextBlob	TextBlob is a Python (2 and 3) library for processing textual data. It provides a simple API for diving into common natural language processing (NLP) tasks such as part-of-speech tagging, noun phrase extraction, sentiment analysis, classification, translation, and more.
Tweets	a message, image, etc. posted on Twitter.
NLP	Natural language processing (NLP) is the ability of a computer program to understand human language as it is spoken.
Cluster Analysis	Cluster analysis is a statistical classification technique in which a set of objects or points with similar characteristics are grouped together
Naïve Bayes Algorithm:	Naive Bayes is a simple technique for constructing classifiers: models that assign class labels to problem instances, represented as vectors of feature values, where the class labels are drawn from some finite set

2. Problem Definition

2.1 Scope

Data Mining is a fairly new concept. a lot of work has been going on in this field. The ability to exploit public sentiment in social media is increasingly considered as an important tool for market understanding, customer segmentation and stock price prediction for strategic marketing planning and maneuvering. This evolution of technology adoption is energized by the healthy growth in big data framework, which caused applications based on Sentiment Analysis (SA) in big data to become common for businesses.

As mentioned before, the main aim of this project is to combine the outputs of various papers under study to try and understand the accuracy of already present algorithms.

Detecting the mood of a person by taking his/her writing as an input (Either a user input or by fetching a file which contains the writing); this may be a social media post or an update on someone's wall. Information gained from applying SA to social media data has many potential usages, for instance, to help marketers evaluate the success of an ad campaign, to identify how different demographics have received a product release, to predict user behavior, or to forecast election results

2.2 Assumptions

- Sentiment analysis is extremely useful in social media monitoring as it allows us to gain an overview of the wider public opinion behind certain topics. The applications of sentiment analysis are broad and powerful. The ability to extract insights from social data is a practice that is being widely adopted by organizations across the world.

- In this project, what could not be captured are sentiments such as sarcasm and bipolar moods.

- Another thing in consideration is that the inputs that are being taken are in English only. Ant input in the form of some other language might not match the words in the database.

3. Project Planning

3.1 Software Lifecycle Model

We are going to follow the Iterative Waterfall Model for our Project.

An iterative waterfall life cycle model does not attempt to start with a full specification of requirements. Instead, development begins by specifying and implementing just part of the software, which can then be reviewed in order to identify further requirements. This process is then repeated, producing a new version of the software for each cycle of the model.

Consider an iterative lifecycle model which consists of repeating the following four phases in sequence:

- A Requirements phase, in which the requirements for the software are gathered and analyzed. Iteration should eventually result in a requirement phase that produces a complete and final specification of requirements.
- A Design phase, in which a software solution to meet the requirements is designed. This may be a new design, or an extension of an earlier design.
- A Design phase, in which a software solution to meet the requirements is designed. This may be a new design, or an extension of an earlier design.
- An Implementation and Test phase, when the software is coded, integrated and tested.
- A Review phase, in which the software is evaluated, the current requirements are reviewed, and changes and additions to requirements.

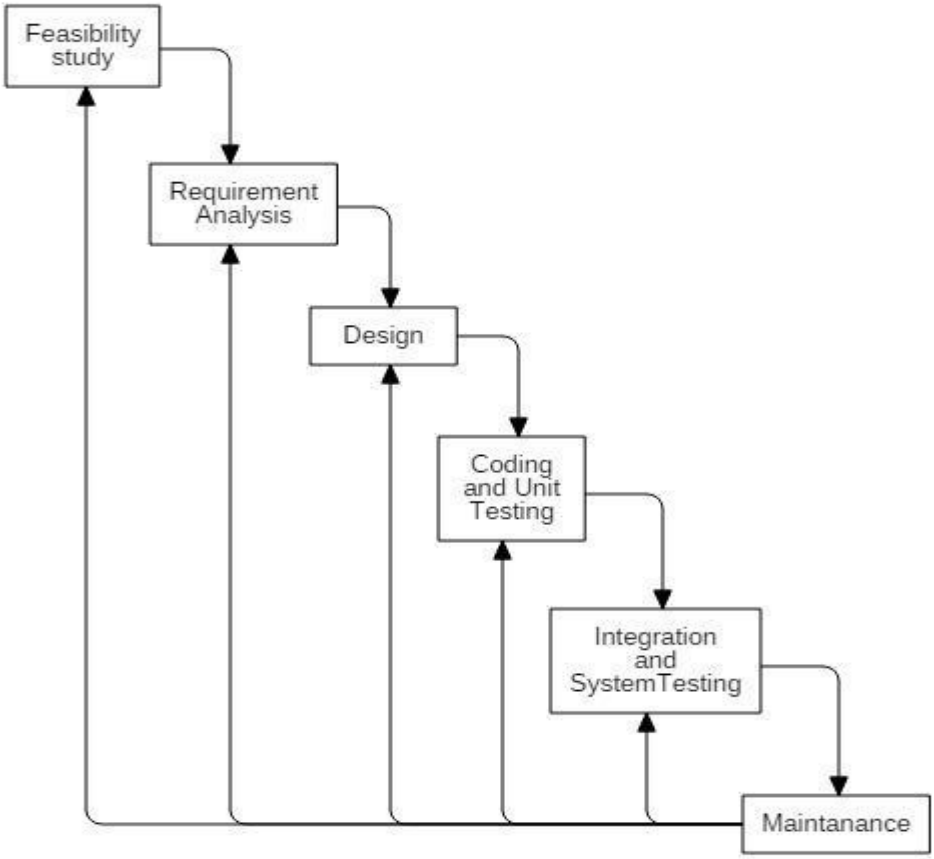


Figure 1 : **Iterative Waterfall Model**

3.2 Scheduling

GANTT CHART

	Task Name	Duration	Start	Finish	%Com	Predeces	Resource Names
1	**Mood Analysis In Social Media**	**174 days**	**Wed 20-09-17**	**Tue 15-05-18**	**100%**		**Tamesh,Priya,Sucharita,Rimpi**
2	**Phase 1: 7th sem Activities**	**54 days**	**Wed 20-09-17**	**Mon 04-12-17**	**100%**		
33	**Semester Break**	**24 days**	**Tue 05-12-17**	**Fri 05-01-18**	**100%**	**2**	
34	**Phase 2: 8th sem Activities**	**75 days**	**Tue 06-02-18**	**Tue 15-05-18**	**100%**	**33**	
35	**Update Requirement analysis**	**58 days**	**Tue 06-02-18**	**Wed 25-04-18**	**100%**		
36	Update requirement matrix	58 days	Tue 06-02-18	Wed 25-04-18	100%		Priya,Rimpi
37	Finalize Requirement matrix	0 days	Wed 25-04-18	Wed 25-04-18	100%	36	Priya,Rimpi,Sucharita,Tamesh
38	**Update Design**	**63 days**	**Thu 22-02-18**	**Tue 15-05-18**	**100%**	**35**	
39	**Detailed Design**	**5 days**	**Tue 17-04-18**	**Mon 23-04-18**	**100%**		
40	Update Use Case Diagram	4 days	Tue 17-04-18	Fri 20-04-18	100%		Tamesh,Priya
41	Update Class Diagram	2 days	Tue 17-04-18	Wed 18-04-18	100%		Rimpi
42	Update Activity Diagram	3 days	Thu 19-04-18	Mon 23-04-18	100%		Rimpi,Sucharita
43	**Test Plan**	**43 days**	**Thu 01-03-18**	**Fri 27-04-18**	**100%**	**39**	
44	Login into Twitter account	1 day	Tue 06-03-18	Tue 06-03-18	100%		Tamesh,Priya
45	Login into TwitterforDevelopers	1 day	Tue 06-03-18	Tue 06-03-18	100%		Tamesh,Priya
46	Verify Consumer Key/ Consumer Secret/ Acc	4 days	Thu 01-03-18	Tue 06-03-18	100%		Tamesh,Priya
47	Sentiment analysis	42 days	Thu 01-03-18	Fri 27-04-18	100%		Tamesh,Sucharita,Rimpi,Priya
48	**Coding & Únit Testing**	**41 days**	**Thu 22-02-18**	**Wed 18-04-18**	**100%**		
49	Coding	36 days	Thu 22-02-18	Thu 12-04-18	100%		Tamesh,Priya,Rimpi,Sucharita
50	Unit Testing	4 days	Sat 14-04-18	Wed 18-04-18	100%		Sucharita
51	**System Integration Testing**	**7 days**	**Wed 18-04-18**	**Thu 26-04-18**	**100%**	**48**	
52	Overall Sentiment analysis	7 days	Wed 18-04-18	Thu 26-04-18	100%		Tamesh,Rimpi,Priya,Sucharita
53	**Project Closure**	**51 days**	**Mon 12-03-18**	**Tue 15-05-18**	**100%**	**51**	
54	Prepare 8th Semester Project Report	44 days	Wed 14-03-18	Wed 09-05-18	100%		Priya,Rimpi
55	Updated Requirement Matrix	28 days	Mon 12-03-18	Tue 17-04-18	100%		Sucharita
56	Updated Project Plan	47 days	Mon 12-03-18	Thu 10-05-18	100%	55	Priya,Rimpi
57	Review by Faculties	0 days	Tue 15-05-18	Tue 15-05-18	100%	56	
58	Approved Project Report - 8th Semester	0 days	Tue 15-05-18	Tue 15-05-18	100%	57	

Figure 2 **: Gantt Chart - Part 1**

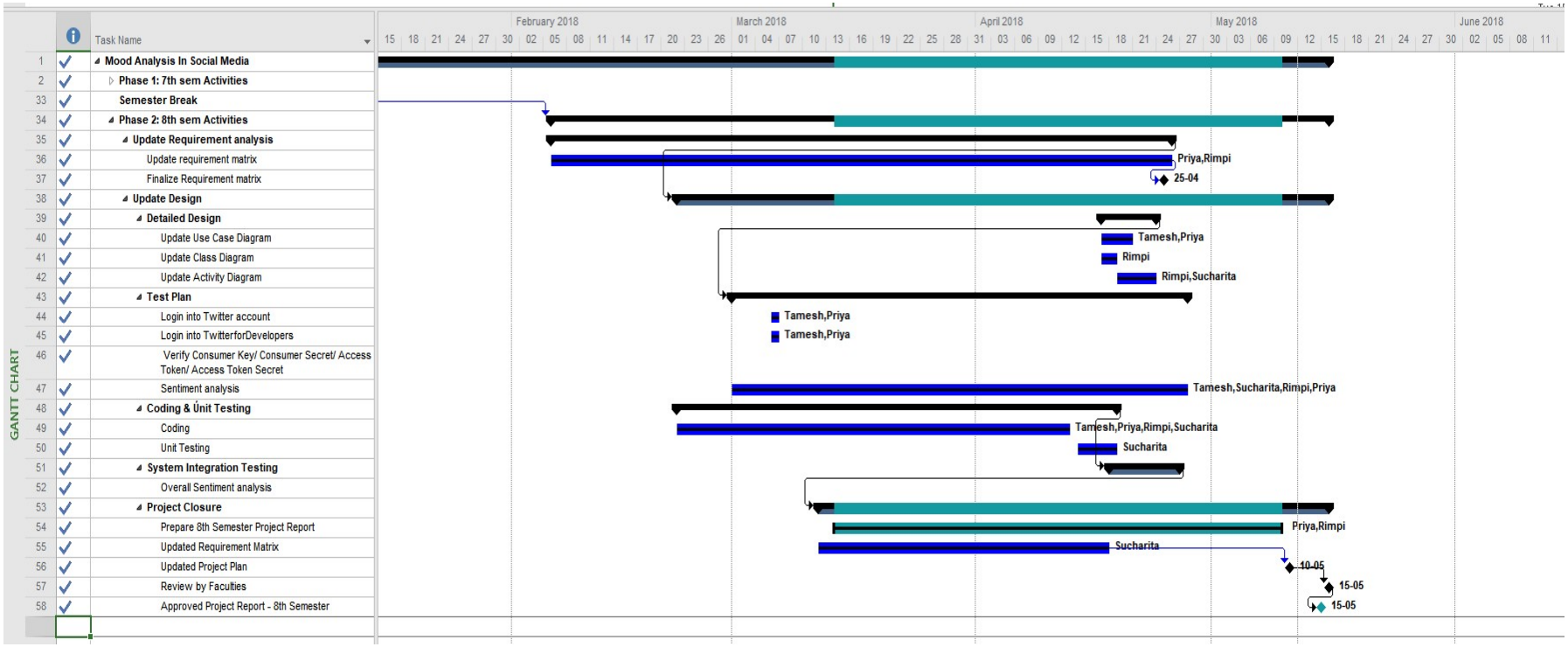


Figure 3 : **Gantt Chart - Part 2**

3.3 Cost Analysis

A project is feasible if it is possible to finish it within fixed budget constraints. Cost Analysis, sometimes also referred to as Cost/Benefit Analysis (CBA), is a systematic approach used to calculate and compare the benefits and costs of a project. For the purpose of this project, we are using a heuristic estimation technique called Constructive Cost Estimation Model (COCOMO).

COCOMO

The basic COCOMO gives an approximate estimate of the project parameters. This estimation model is given by the following expressions:

$$Effort\ Applied\ (E) = a_b \times (KLOC)^{b_b}\ PM$$

$$Development\ Time\ (D) = c_b \times (Effort\ Applied)^{d_b}\ months$$

Where:

i. KLOC is the estimated size of the software product expressed in Kilo Lines of Code.
ii. Effort Applied is the total effort required to develop the software product, expressed in person-months (PMs).
iii. Development Time is the estimated time to develop the software, expressed in months.
iv. a_b, b_b, c_b and d_b are constants as given in the following table:

Estimation of effort, development time and productivity

Table 2 : ***COCOMO Model***

Software Project	a_b	b_b	c_b	d_b
Organic	2.4	1.05	2.5	0.38
Semi-detached	3.0	1.12	2.5	0.35
Embedded	3.6	1.20	2.5	0.32

** We use the Organic COCOMO Model here.

Lines of code = 155 = .155 KLOC (estimated)

Effort = $2.4 \times (.155)^{1.05}$ = .34 man-months (estimated)

Development Time = $2.5 \times (.34)^{0.38}$ = 1.66 months

Productivity = Effort/Development time
= .34/1.66 man-months/month

4. Requirement Analysis

4.1 Requirement Matrix

Table 3: ***Requirement Matrix: Traceability Matrix***

Rqmt ID	*Requirement Item*	*Requirement Status*	*Design Module*	*Design Reference (section# under project Report)*	Test Case Number	*Technical Platform of Implementation*	*Prototype prepared ?*	Name of Program / Component	Test Results Reference	Additional Comments (if not included in previous columns)
TW	Login into Twitter Account	Completed	TW	4.2.1	T-TW-1,2,3	Twitter API	Yes	Browser	T-TW-1,2,3	Server Site
TD-1	Login into TwitterForDevelopers	Completed	TD	4.2.2	T-TD-1,2	Twitter API	Yes	Browser	T-TD-1,2	Server Site
TD-1.1	Create new App	Completed	TD	4.2.2.1	T-TD-3	Twitter API	Yes	Browser	T-TD-3	Server Site
TD-1.2	Get Consumer Key/ APIKey	Completed	TD	4.2.2.2	T-TD-1,2	Twitter API	Yes	Browser	T-TD-1,2	Server Site
TD-1.3	Get Consumer Secret/ API Secret	Completed	TD	4.2.2.3	T-TD-1,2	Twitter API	Yes	Browser	T-TD-1,2	Server Site
TD-1.4	Get Access Token	Completed	TD	4.2.2.4	T-TD-1,2	Twitter API	Yes	Browser	T-TD-1,2	Server Site
TD-1.5	Get Access Token Secret	Completed	TD	4.2.2.5	T-TD-1,2	Twitter API	Yes	Browser	T-TD-1,2	Server Site
VER-1	Verify API Key	Completed	VER	4.2.3	T-VER-1,2	Python	Yes	MoodAnalysis.py	T-VER-1,2	

VER-1.1	Verify API Secret	Completed	VER	4.2.3.1	T-VER-1,2	Python	Yes	MoodAnalysis.py	T-VER-1,2	
VER-1.2	Verify Access Token	Completed	VER	4.2.3.2	T-VER-1,2	Python	Yes	MoodAnalysis.py	T-VER-1,2	
VER-1.3	Verify Access Token Secret	Completed	VER	4.2.3.3	T-VER-1,2	Python	Yes	MoodAnalysis.py	T-VER-1,2	
MA	Sentiment Analysis	Completed	MA	4.2.4	T-MA-1,2,3,4,5,6	Python	Yes	MoodAnalysis.py	T-MA-1,2,3,4,5,6	

4.2 Requirement Elaboration

4.2.1 TW:
Login into the Twitter Account.

4.2.2 TD-1:
Login into Twitter for Developers using the same account credentials

4.2.2.1 TD-1.1:
Click on Create Application button under Details menu

4.2.2.2 TD-1.2:
Get Consumer Key/ App Key

4.2.2.3 TD-1.3:
Get Consumer Secret/ App Secret

4.2.2.4 TD-1.4:
Get Access Token

4.2.2.5 TD-1.5:
Get Access Token Secret

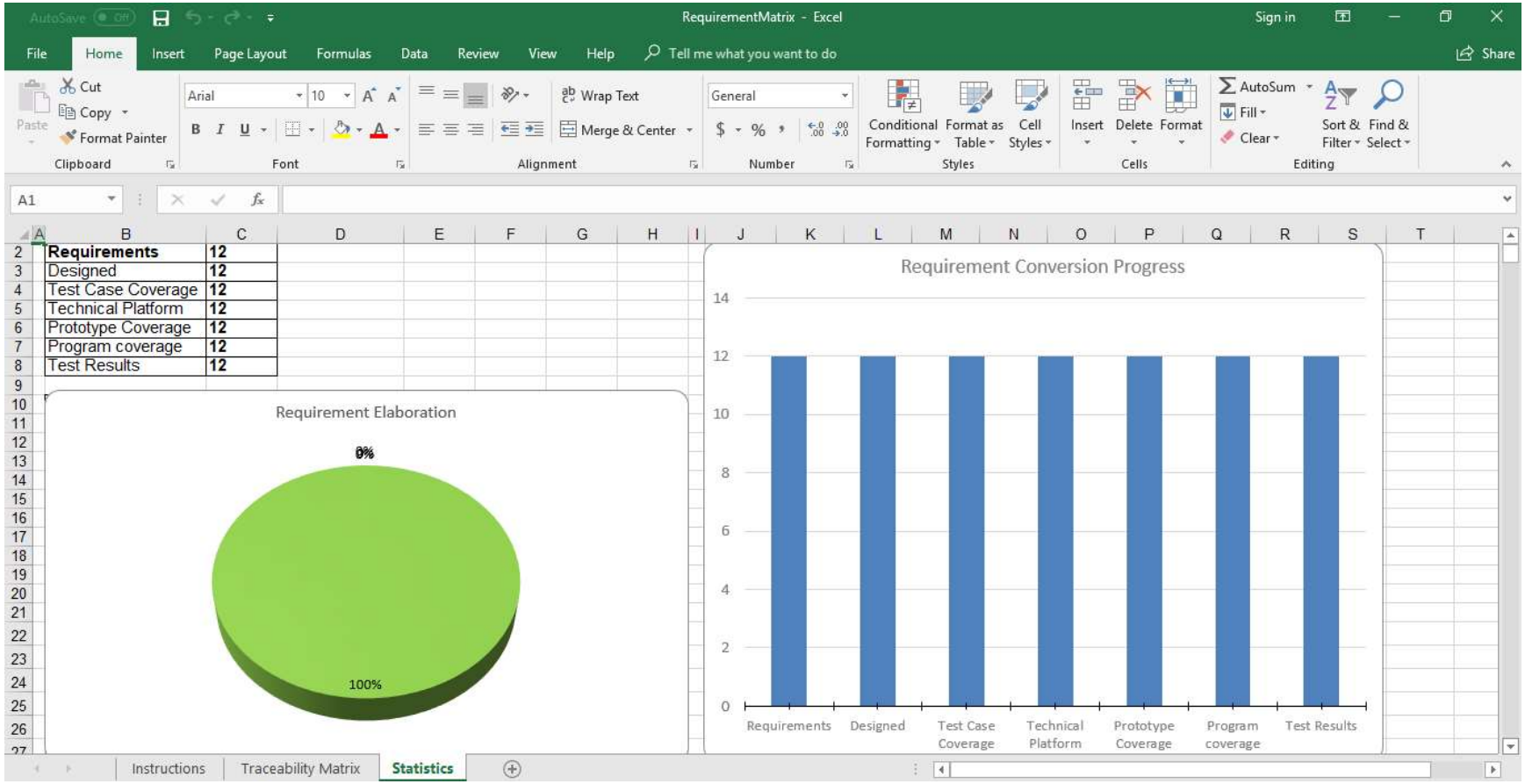
4.2.3 VER-1:
Verify Consumer Key/ App ID.

4.2.3.1 VER-1.1:
Verify Consumer Secret/ App Secret

4.2.3.2 VER-1.2:
Verify Access Token

4.2.3.3 VER-1.3:
Verify Access Token Secret

4.2.4 MA:
Run the Code for Sentiment Analysis



Requirements	**12**
Designed	**12**
Test Case Coverage	**12**
Technical Platform	**12**
Prototype Coverage	**12**
Program coverage	**12**
Test Results	**12**

Figure 4 : **Requirement Elaboration**

5. Design

5.1 Technical Environment

Python 3.5.2

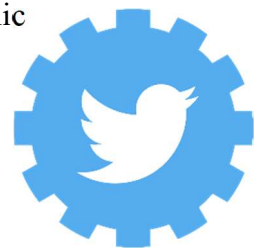
Python is a dynamic object-oriented programming language that can be used for many kinds of software development. It offers strong support for integration with other languages and tools, comes with extensive standard libraries, and can be learned in a few days. Many Python programmers report substantial productivity gains and feel the language encourages the development of higher quality, more maintainable code.

Python runs on Windows, Linux/Unix, Mac OS X, OS/2, Amiga, Palm Handhelds, and Nokia mobile phones. Python has also been ported to the Java and .NET virtual machines.

Python is distributed under an OSI-approved open source license that makes it free to use, even for commercial products.

Twitter for Developers

Twitter data is the most comprehensive source of live, public conversation worldwide. The REST, streaming, and Enterprise APIs enable programmatic analysis of Tweets back to the first Tweet in 2006. Whether someone is building a solution for brands, team, or conducting research—the endpoints enable rich insight into audiences, market movements, emerging trends, key topics, breaking news, and much more.



Example use cases: Trend detection, audience insights, and research.

5.2 Hierarchy of Modules

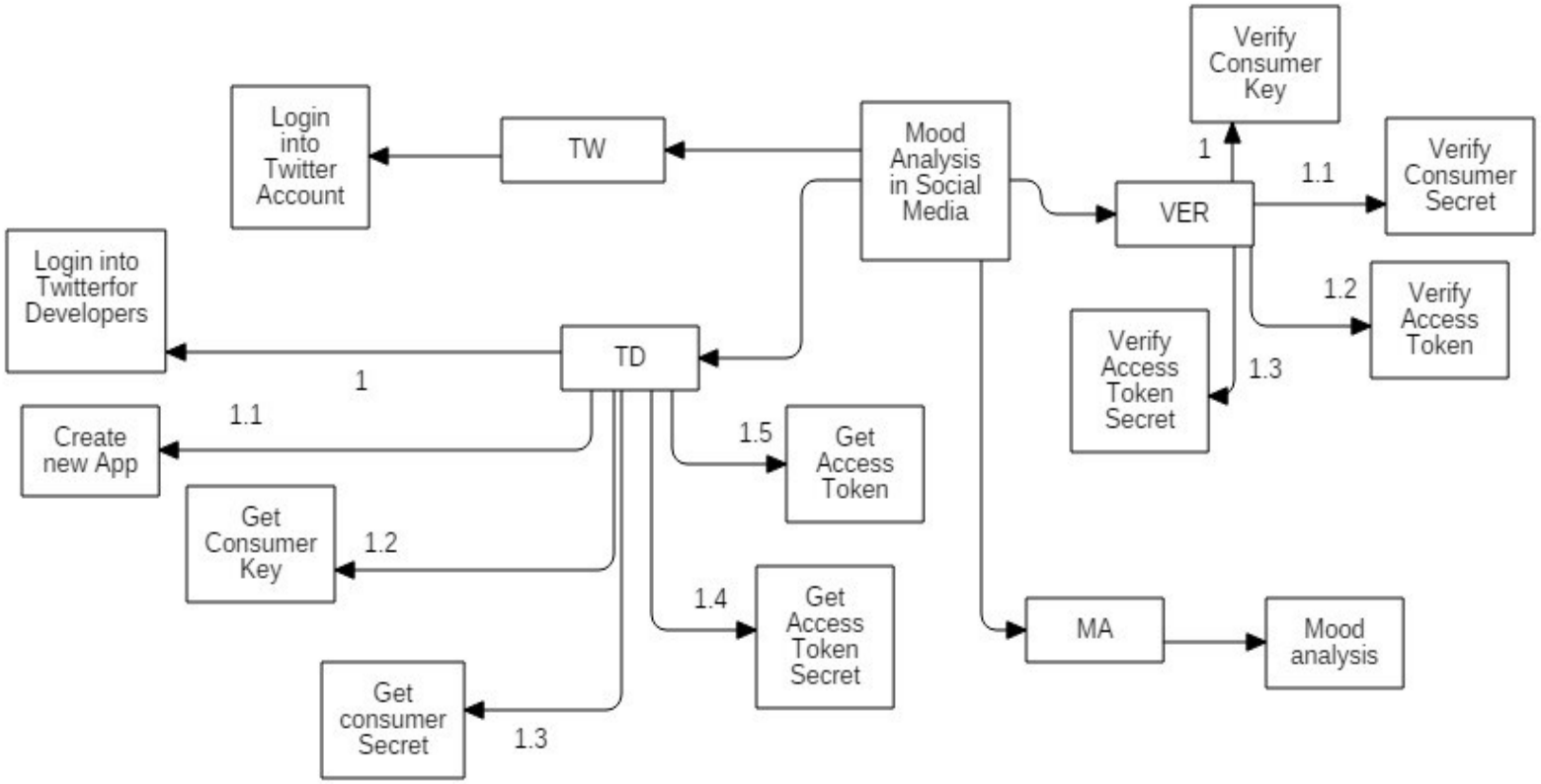


Figure 5 **: Hierarchy of Modules**

5.3 Detailed Design

5.3.1 Use Case Diagram

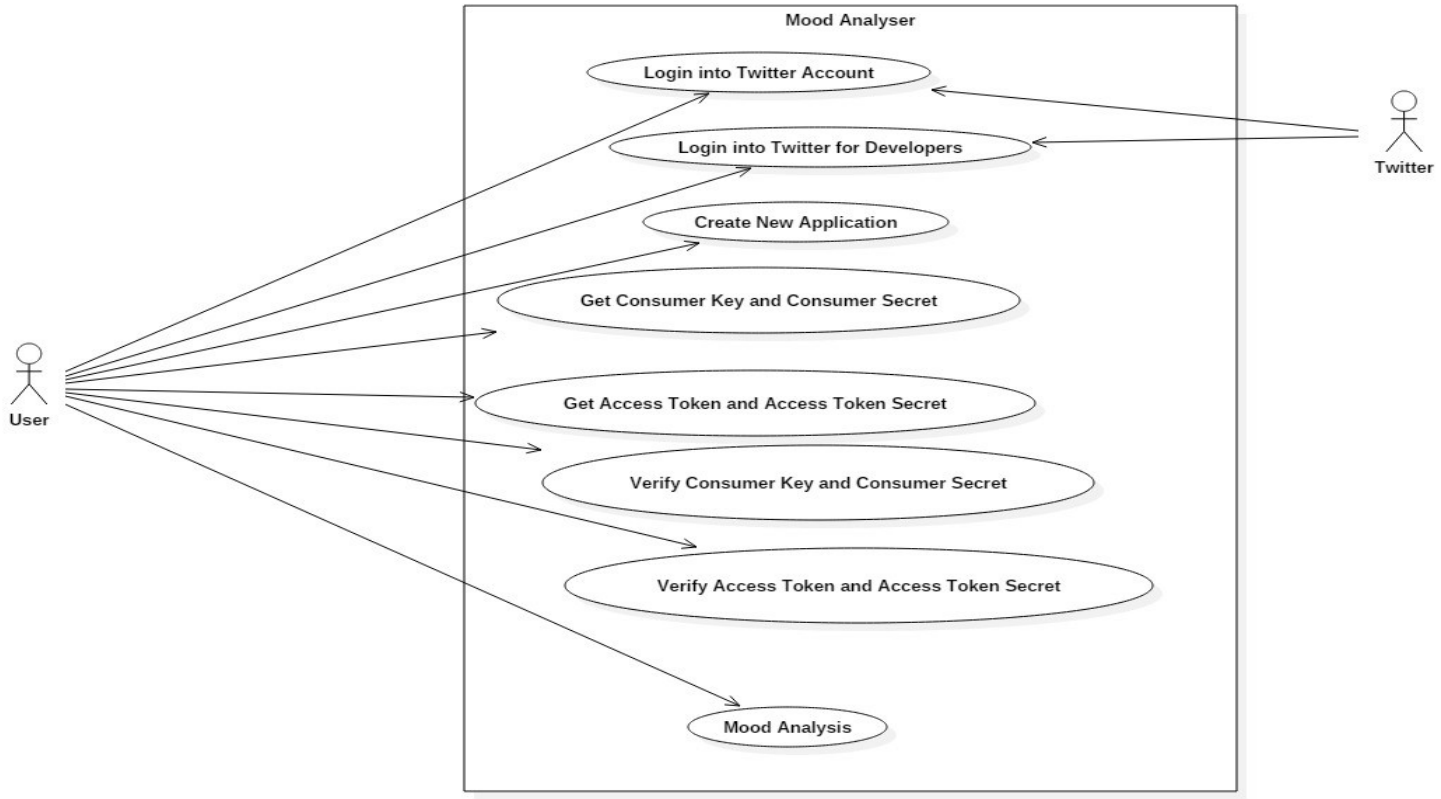


Figure 6 : **Use Case Diagram**

5.3.2 Class Diagram

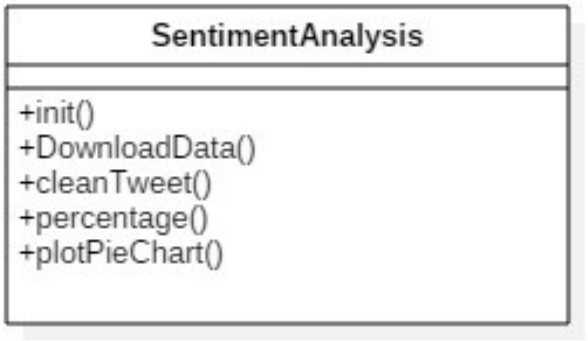


Figure 7 : **Class Diagram for Class SentimentAnalysis**

5.3.3 Activity Diagram

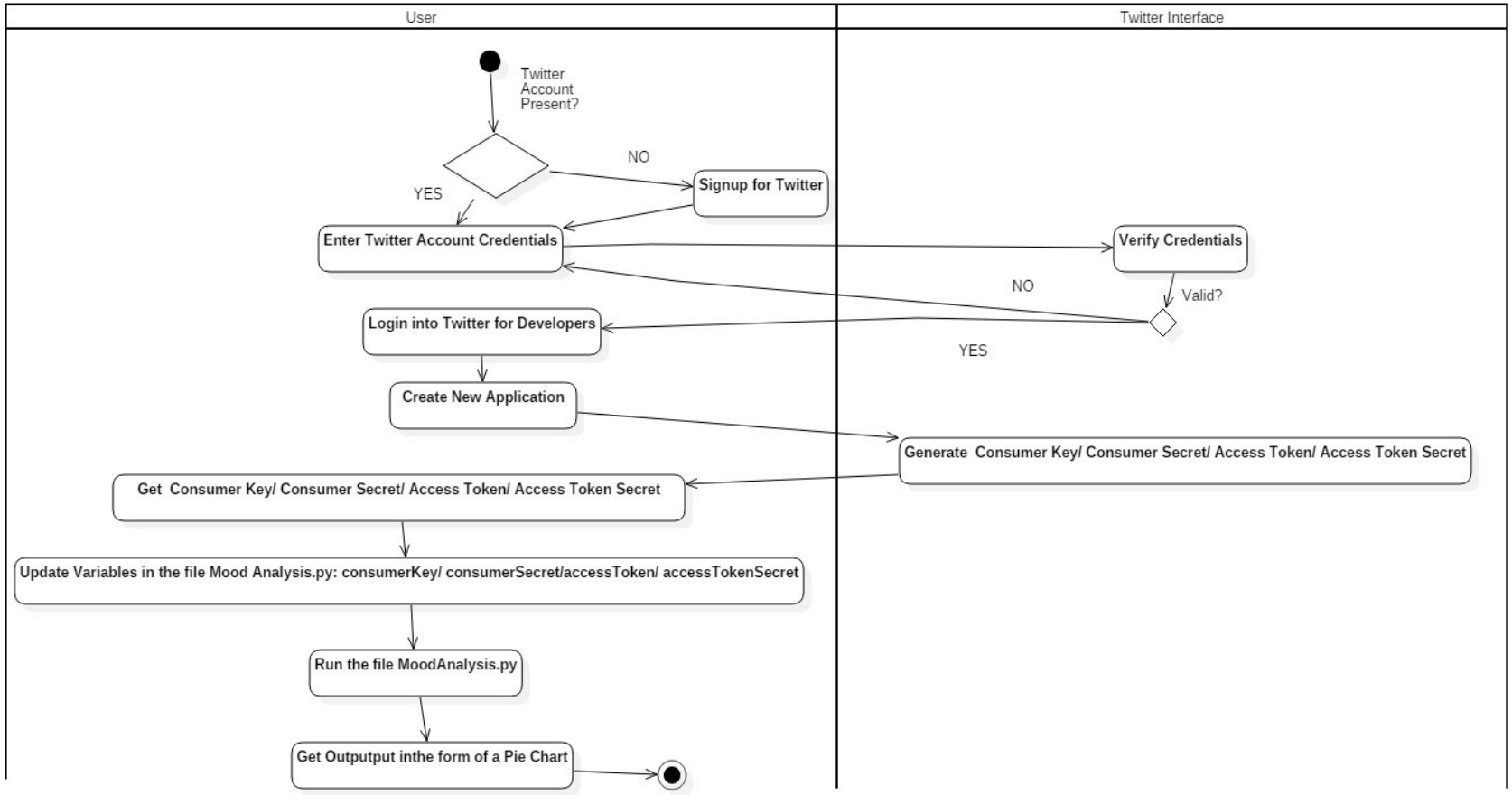


Figure 8 **: Activity Diagram**

5.4 Test Plan

*Table 4 : **Test Planning***

Test Case Number	Test Case	Input	Expected Result	Observed Result	Status
T-TW-1	Wrong Login Credentials	Enter Wrong Credentials	The email and password you entered did not match our records. Please double-check and try again.	The email and password you entered did not match our records. Please double-check and try again.	O.K.
T-TW-2	Correct Login Credentials	Enter correct credentials	Profile Home Page opens	Profile Home Page opens	O.K.
T-TW-3	No Twitter account	Incase developer does not have a Facebook page.	Access Denied	Access Denied	O.K.
T-TD-1	Wrong Login Credentials	Enter Wrong Credentials	The email and password you entered did not match our records. Please double-check and try again.	The email and password you entered did not match our records. Please double-check and try again.	O.K.
T-TD-2	Correct Login Credentials	Enter correct credentials	Access Granted and developer's profile opens.	Access Granted and developer's profile opens.	O.K.

T-TD-3	Create New Application	Click on Create New Application Button under Details	Access Token/ Access Token Secret is generated	Access Token/ Access Token Secret is generated	OK
T-VER-1	Wrong App Secret/App ID/ Access Token/ Access Token Secret	Enter wrong App Secret/App ID/ Token Access/ Access Token Secret	Enter Keyword/Tag to search about:	Enter Keyword/Tag to search about:	OK
T-VER-2	Correct App Secret/App ID/ Access Token/ Access Token Secret	Enter correct App Secret/App ID/ Access Token/ Access Token Secret	Enter Keyword/ Tag to search about:	Enter Keyword/ Tag to search about:	OK
T-MA-1	Enter Keyword/ Tag to search about:	Arsenal	Enter number of tweets to search:	Enter number of tweets to search:	OK
T-MA-2	Enter number of tweets to search in negative integer	-456	Invalid Input. Try again. Enter how many tweets to search:	Invalid Input. Try again. Enter how many tweets to search:	OK
T-MA-3	Enter 0 as number of tweets to search	0	Invalid Input. Try again. Enter how many tweets to search:	Invalid Input. Try again. Enter how many tweets to search:	OK

T-MA-4	Enter number of tweets to search in positive integer	100	The extraction of feed starts, and pie chart is displayed with the result	The extraction of feed starts, and pie chart is displayed with the result	OK
T-MA-5	Enter Number of tweets to search as an alphabet/ character	a	Enter Keyword/Tag to search about:	Enter Keyword/Tag to search about:	OK
T-MA-6	Enter Number of tweets to search as an alphabet/ character	%	Enter Keyword/Tag to search about:	Enter Keyword/Tag to search about:	OK

Table 5 : ***Test Results***

Test Case Number	Test Reference Number	Figure Reference
T-TW-1	T-TW-1	Figure 9
T-TW-2	T-TW-2	Figure 10
T-TW-3	T-TW-3	Figure 11
T-TD-1	T-TD-1	Figure 12,13
T-TD-2	T-TD-2	Figure 12, 14
T-TD-3	T-TD-3	Figure 15
T-VER-1	T-VER-1	Figure 16
T-VER-2	T-VER-2	Figure 17
T-MA-1	T-MA-1	Figure 17
T-MA-2	T-MA-2	Figure 17
T-MA-3	T-MA-3	Figure 18
T-MA-4	T-MA-4	Figure 19
T-MA-5	T-MA-5	Figure 20
T-MA-6	T-MA-6	Figure 21

6. Implementation

6.1 Implementation Details

A naive Bayes classifier is an algorithm that uses Bayes' theorem to classify objects. Naive Bayes classifiers assume strong, or naive, independence between attributes of data points. Popular uses of naive Bayes classifiers include spam filters, text analysis and medical diagnosis. These classifiers are widely used for machine learning because they are simple to implement.

4.2.1 TW:
Login into the Twitter Account.

4.2.2 TD-1:
Login into Twitter for Developers using the same account credentials

4.2.2.1 TD-1.1:
Click on Create Application under Details menu

4.2.2.2 TD-1.2:
Get Consumer Key/ App Key

4.2.2.3 TD-1.3
Get Consumer Secret/ App Secret

4.2.2.4 TD-1.4
Get Access Token

4.2.2.5 TD-1.5
Get Access Token Secret

4.2.3 VER-1:
Verify Consumer Key/ App ID.

4.2.3.1 VER-1.1:
Verify Consumer Secret/ App Secret

4.2.3.2 VER-1.2:
Verify Access Token

4.2.3.3 VER-1.3
Verify Access Token Secret

4.2.4 MA
Run the Code for Sentiment Analysis

6.2 System Installation Steps

1. Install Python 3.5.2 in your device
2. Set path variable accordingly
3. Run command prompt and run the following command:
    i. pip install -- upgrade pip
    ii. pip install textblob tweepy matplotlib

6.3 System Usage Information

1. Create a twitter account
2. Open Twitter for Developers using the same account
3. Fill in the necessary details in the form
4. Select the Create Application option under Details
5. Go to Keys and Access Tokens to copy and paste the Consumer Key, Consumer Secret, Access Token and Access Token Secret in the respective place in the code in MoodAnalysis.py
6. Run command prompt
7. Set path according to the location of the **MoodAnalysis.py** file
8. Run the following command:
   py MoodAnalysis.py
9. Enter the Keyword/Tag you want to search
10. Enter the number of tweets you want to search

7. Conclusion

7.1 Project Benefits

We draw certain insights from the research papers that we have thoroughly studied for this purpose. After going through the above mentioned, we have noticed certain flaws i.e. Researches have restricted themselves to basic three supersets of mood i.e. positive, negative and neutral, whilst there is ample room for including more emotions for a thorough perspective on human behavior.

Thus, we think that if we can take the subsets of these supersets into account then the level of accuracy in predicting the mood of a person will increase and scale them with color codes.

As mentioned before, implications of status updates and hashtags holds a lot of importance in the analysis of the mood in social media users. Companies can use this to gauge the response on their products or services which means this has a viable and commercial future scope.

7.2 Future Scope for Improvements

Data mining in social media is a concept that has not been commercialized to its optimum potential and although a lot of research work has already been done, there has not been any concrete analysis or results pertaining to mood assessment and forecasting of social media users. We have undertaken this project for the above-mentioned reason and because we believe this is not a thoroughly explored domain.

As per our understanding, the following mentioned points can be worked on in the near future: -

- ❖ Consideration of “sarcasm” and “irony”
- ❖ Developing fully automatic analyzing tool
- ❖ Effective Analysis of policy opinionated content
- ❖ Successful handling of bipolar sentiments
- ❖ Generation of highly content lexicon database
- ❖ Develop models that can take up inputs in various languages

8. References

- [1] Twitter Sentiment Analysis
https://dl.acm.org/citation.cfm?id=1944571

- [2] Sentiment Analysis: Facebook Status Messages
http://people.sabanciuniv.edu/berrin/share/LDA/Stanford-NLP-Course-termproject-ssoriajr-kanej.pdf

- [3] Recognizing contextual polarity in phrase-level sentiment analysis
https://dl.acm.org/citation.cfm?id=1220619

- [4] Automatically predicting mood from expressed emotions
https://repository.tudelft.nl/islandora/object/uuid%3A12355ac5-e752-43e6-a26c-cb0528c48977

- Twitter
https://twitter.com

- Twitter for Developers
https://apps.twitter.com

- Python Tutorials
https://pythonspot.com

Appendix A (Prototype)

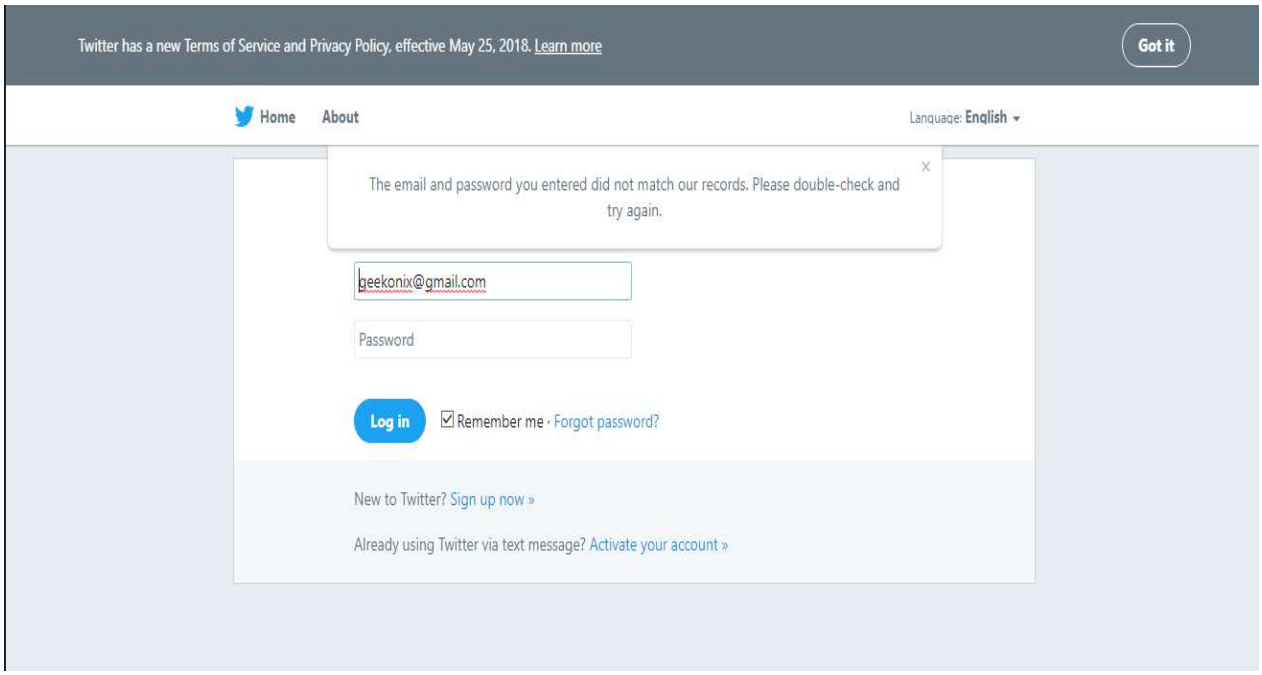


Figure 9: Twitter Login with Wrong Credentials

Figure 10: Twitter profile Home Page

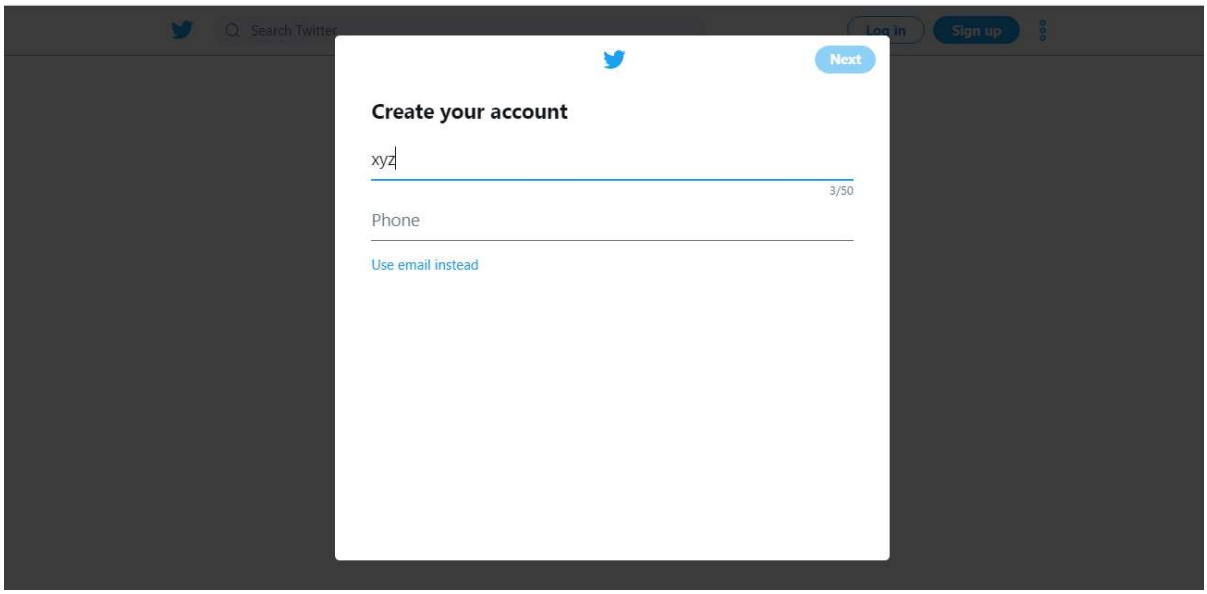


Figure 11: Twitter Sign Up page

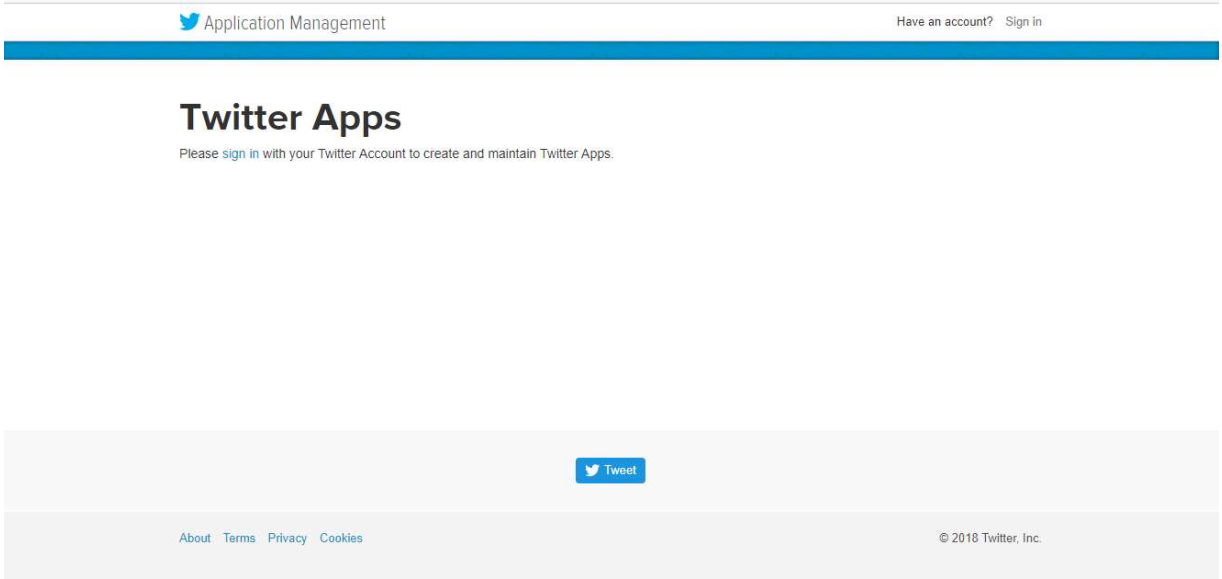


Figure 12: apps.twitter Sign In page

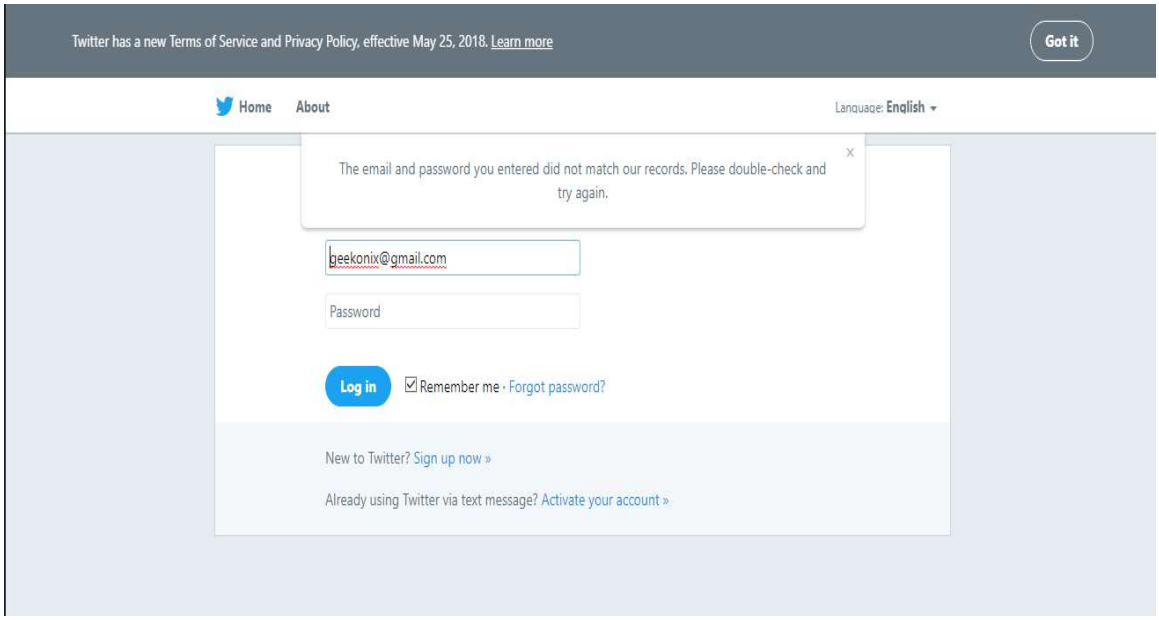


Figure 13: Login into apps.twitter using incorrect credentials

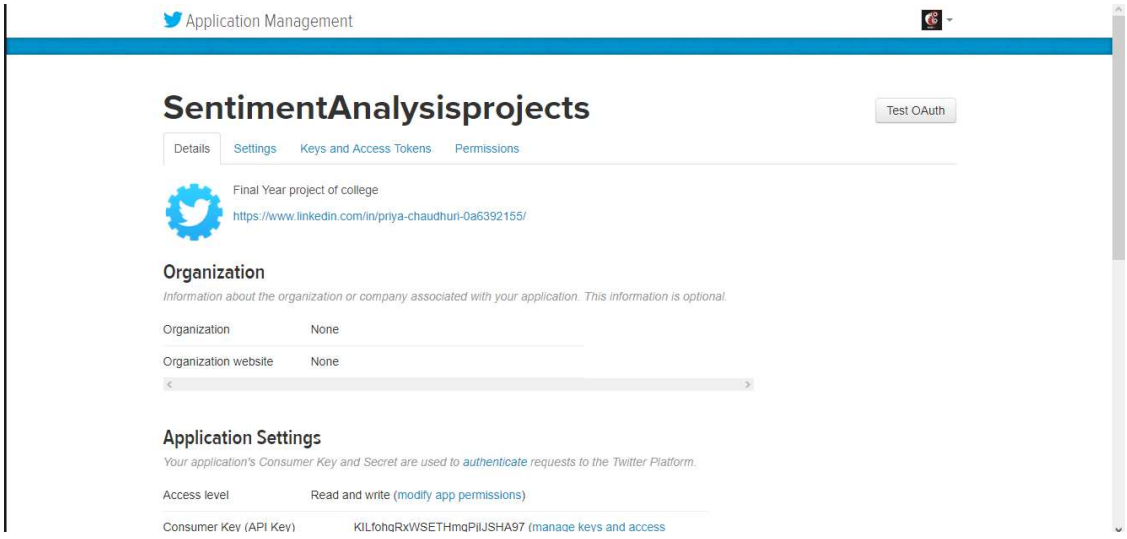


Figure 14: Login into apps.twitter using correct credentials

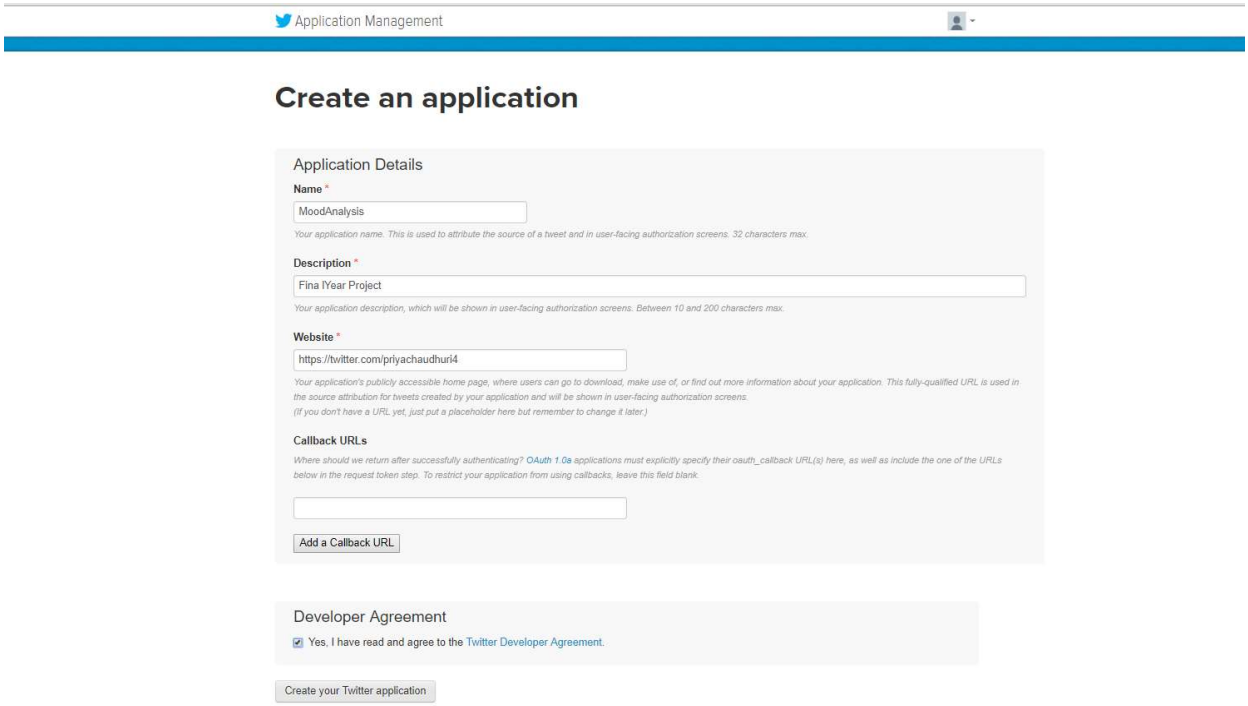


Figure 15: Create New Application

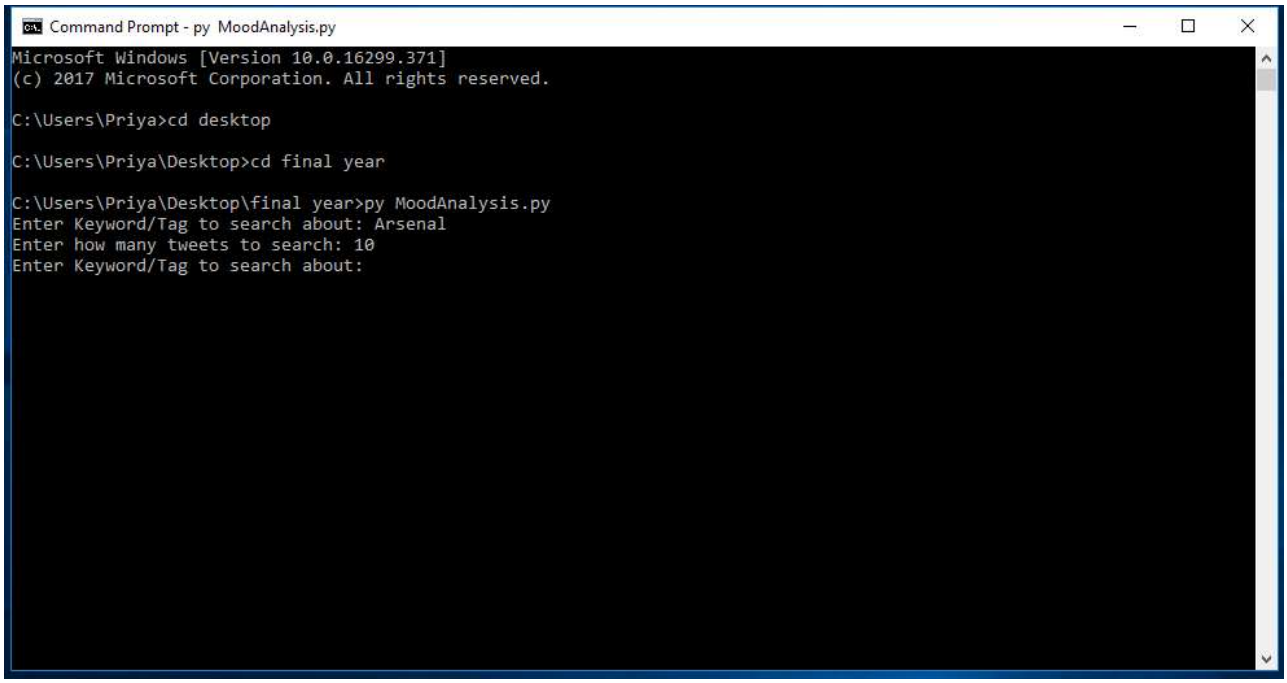


Figure 16: Code run with wrong consumer key

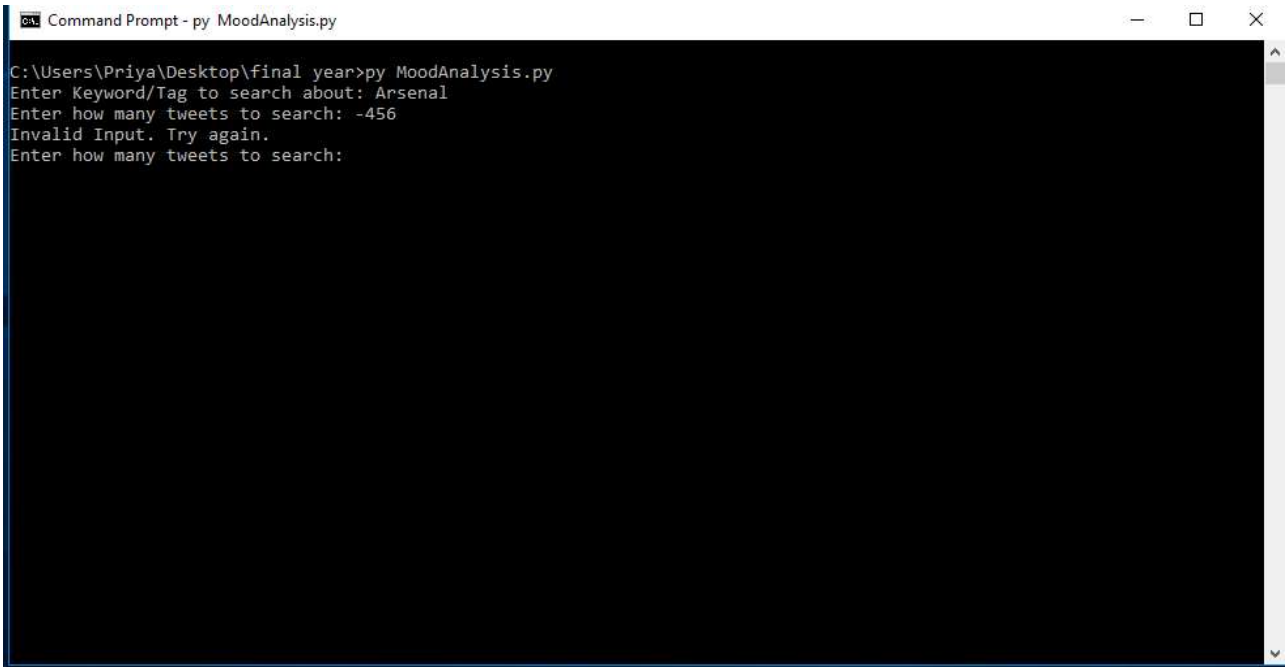


Figure 17 : Code run with correct consumer key but a negative number of tweets to search

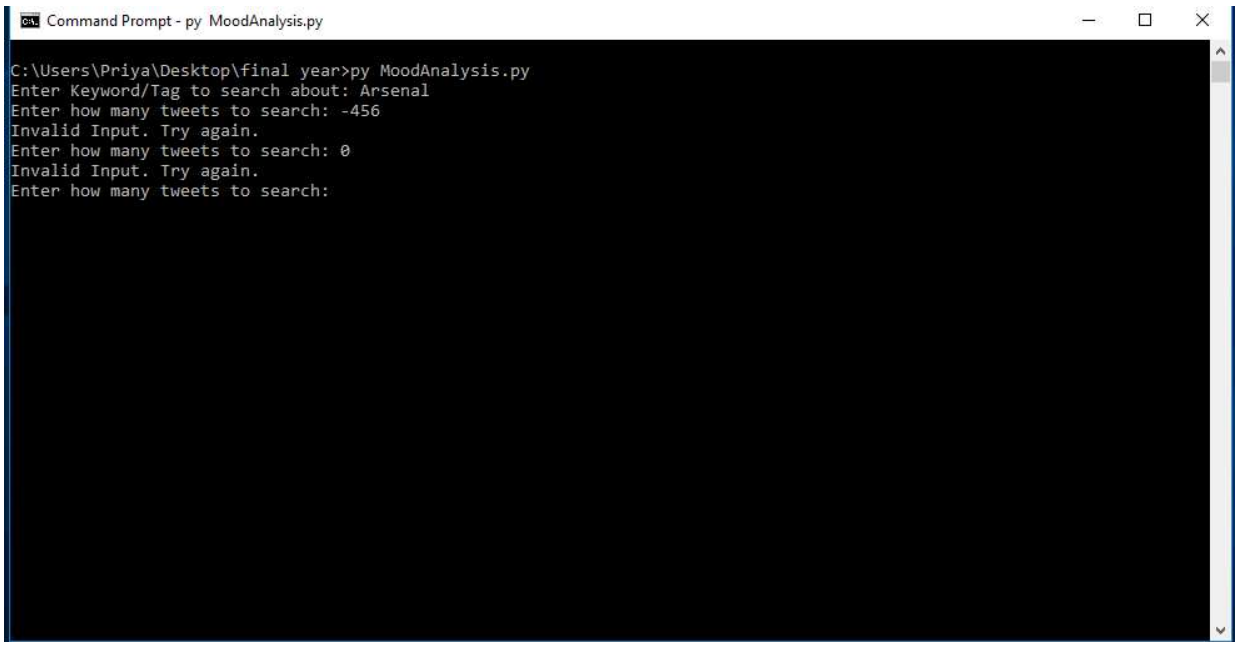


Figure 18: Code run with number of tweets to search as 0

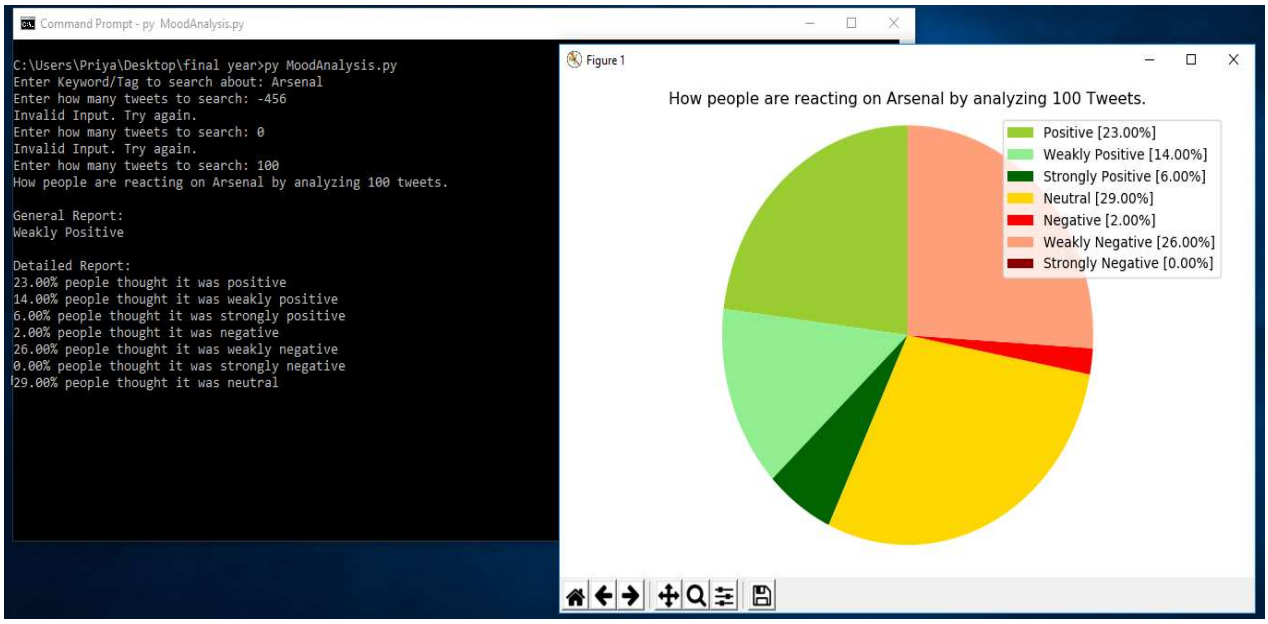


Figure 19: Code run with number of tweets to search as 100

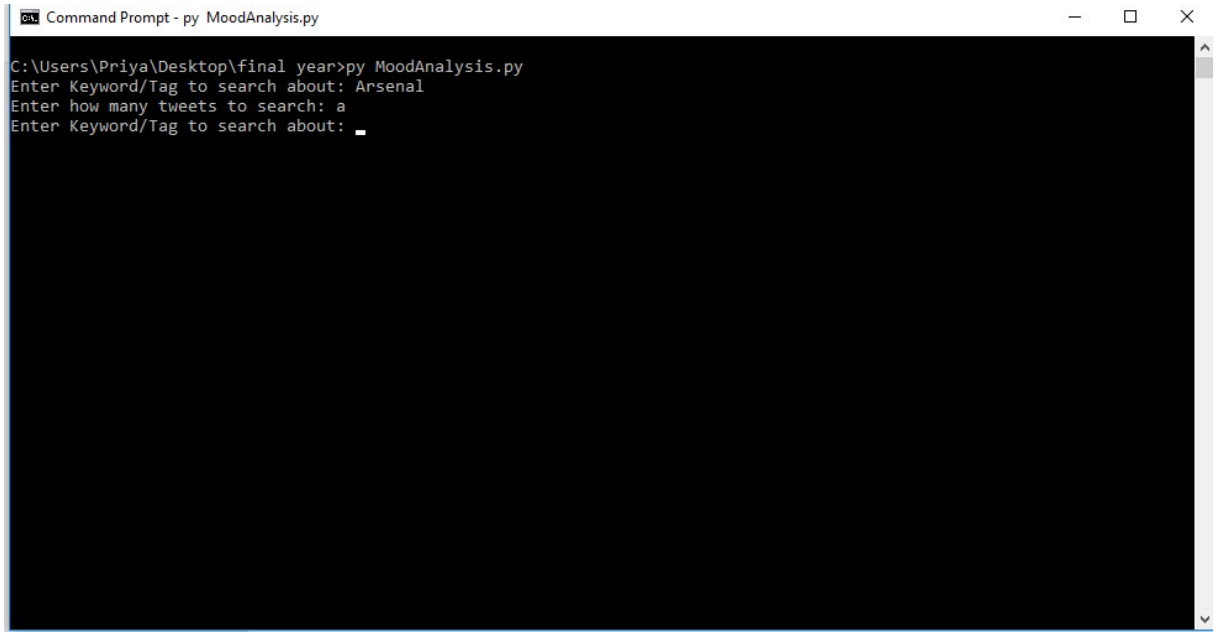


Figure 20: Code run with number of tweets to search as an alphabet

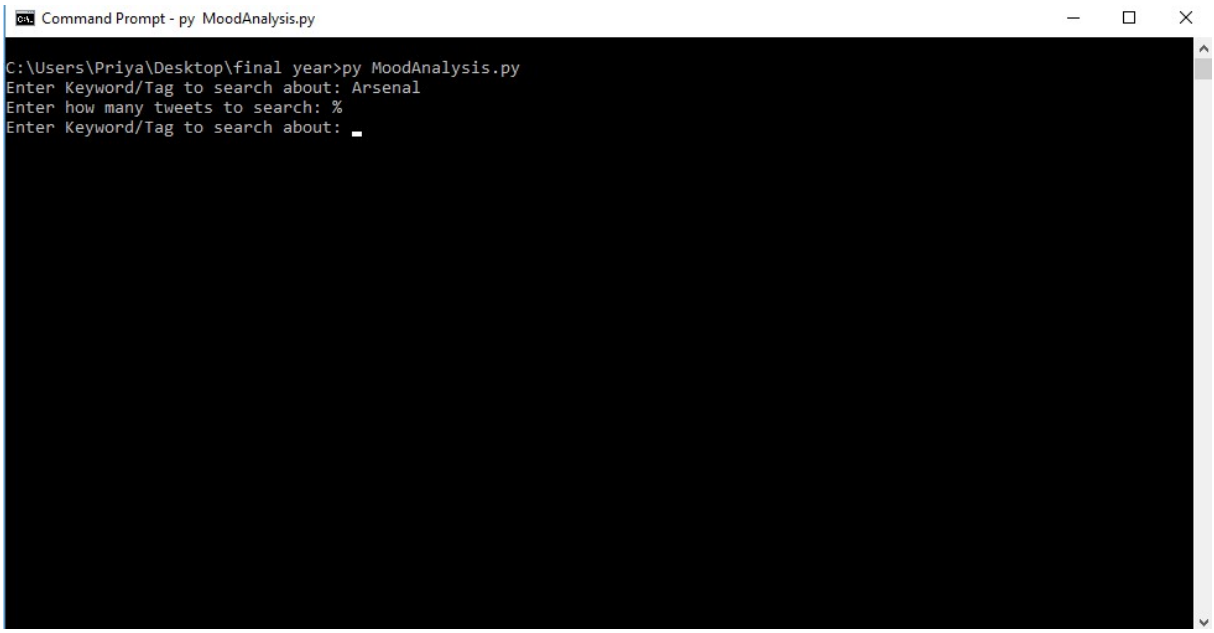


Figure 21: Code run with number of tweets to search as a special character