

Integrated Academy of Management and Technology, Ghaziabad

Presentation/Seminar Based on Minor Project

BCA-505P



Integrated Academy of Management and Technology, Ghaziabad
NH 24, Near Dasna Crossing, Adhyatmik Nagar, Udyog Kunj,
Ghaziabad, Uttar Pradesh 201009

Integrated Academy of Management and Technology, Ghaziabad

CONTENTS OF BCA MINOR PROJECT REPORT

S NO.	CONTENTS	PAGE NO.
1	Acknowledgement.	4
2	Introduction.	7
3	Objective Of The Project.	9
4	Analysis.	10
5	System Software Requirement Specifications (SRS).	11
6	Table and Structure, Number of Modules, Detail of Modules	11
7	Code of Sentiment Analysis	12
8	Scope Of The Project	16

Integrated Academy of Management and Technology, Ghaziabad

NH 24, Near Dasna Crossing, Adhyatmik Nagar, Udyog Kunj,
Ghaziabad, Uttar Pradesh 201009

|

Integrated Academy of Management and Technology, Ghaziabad

MINOR PROJECT

BCA-505P

Data Mining and Sentiment Analysis of Twitter Data

MINOR PROJECT

SUBMITTED TO

Prof. Neelam Yadav

**C.C.S. UNIVERSITY
MEERUT (U.P)**

**FOR THE PARTIAL FULFILLMENT OF THE DEGREE OF BACHELOR
IN COMPUTER APPLICATION SESSION 2018-2021**

Under the Guidance of:

(Prof.) Neelam Yadav

Mentor Designation

Department

INMANTEC

Submitted by:

Abhishek

R180845106003

B.C.A. (2018-20210

INMANTEC

Integrated Academy of Management and Technology, Ghaziabad
NH 24, Near Dasna Crossing, Adhyatmik Nagar, Udyog Kunj,
Ghaziabad, Uttar Pradesh 201009

Integrated Academy of Management and Technology, Ghaziabad

MINOR PROJECT BCA-505P

ACKNOWLEDGEMENT

I am very grateful to my Minor Project Mentor **Mrs. Neelam Yadav** for giving his/her valuable time and constructive guidance in preparing the Minor Project. It would not have been possible to complete this Minor Project in short period of time without his/her kind encouragement and valuable guidance.

Date:

Signature

Abhishek

Integrated Academy of Management and Technology, Ghaziabad

MINOR PROJECT BCA-505P

CERTIFICATE OF ORIGINALITY

I hereby declare that my Minor Project (BCA – 505P) titled “ **Data Mining and Sentiment Analysis of Twitter Data** “ submitted to CCS UNIVERSITY (Meerut U.P.) for the partial fulfillment of the degree of Bachelor In Computer Application Session 2018-2021 from INTEGRATED ACADEMY OF MANAGEMENT AND TECHNOLOGY, GHAZIABAD has not previously formed the basis for the award of any other degree, diploma or other title.

Place:

Date:

Signature

Abhishek

Integrated Academy of Management and Technology, Ghaziabad

NH 24, Near Dasna Crossing, Adhyatmik Nagar, Udyog Kunj,
Ghaziabad, Uttar Pradesh 201009

Integrated Academy of Management and Technology, Ghaziabad

Minor Project Name
Team Members

s.No	Roll No	Name	Work Done
1.	180845106003	Abhishek	All

Integrated Academy of Management and Technology, Ghaziabad

NH 24, Near Dasna Crossing, Adhyatmik Nagar, Udyog Kunj,
Ghaziabad, Uttar Pradesh 201009

Introduction

- **Sentiment analysis** is the automated process of analyzing text data and sorting it into sentiments positive, negative, or neutral.
- Using sentiment analysis tools to analyze opinions in Twitter data can help companies understand how people are talking about their brand.
- With [more than 321 million active users, sending a daily average of 500 million Tweets](#), Twitter allows businesses to reach a broad audience and connect with customers without intermediaries. On the downside, it's harder for brands to quickly detect negative content, and if it goes viral you might end up with an unexpected PR crisis on your hands.
- **Sentiment analysis** (also known as ***opinion mining***) is the automated process of identifying and extracting the subjective information that underlies a text. This can be either an opinion, a judgment, or a feeling about a particular topic or subject. The most common type of sentiment analysis is called 'polarity detection' and involves classifying a statement as 'positive', 'negative', or 'neutral'.
- For example, let's take this sentence: *"I don't find the app useful: it's really slow and constantly crashing"*. A sentiment analysis model would automatically tag this as *Negative*.

Integrated Academy of Management and Technology, Ghaziabad

- By analyzing social media posts, product reviews, customer feedback, and NPS responses (among other unstructured data), businesses can understand how their customers *feel* about their product or service.
- Sentiment analysis is particularly useful for social media monitoring because it goes beyond the number of likes or retweets, by providing qualitative insights.

R and Python (DBMS tools) are widely used for Sentiment Analysis of dataset Twitter.

Integrated Academy of Management and Technology, Ghaziabad

NH 24, Near Dasna Crossing, Adhyatik Nagar, Udyog Kunj,
Ghaziabad, Uttar Pradesh 201009

Objectives

Why Twitter Sentiment Analysis:

Sentiment Analysis Dataset Twitter has a number of applications:

- 1. Business:** Companies use Twitter Sentiment Analysis to develop their business strategies, to assess customers' feelings towards products or brand, how people respond to their campaigns or product launches and also why consumers are not buying certain products.
- 2. Politics:** In politics Sentiment Analysis Dataset Twitter is used to keep track of political views, to detect consistency and inconsistency between statements and actions at the government level. Sentiment Analysis Dataset Twitter is also used for analyzing election results.
- 3. Public Actions:** Twitter Sentiment Analysis also is used for monitoring and analyzing social phenomena, for predicting potentially dangerous situations and determining the general mood of the blogosphere.

Why Is Twitter Sentiment Analysis important?

These are some of the main advantages of Twitter sentiment analysis:-

1. **Real-Time Analysis:** Twitter sentiment analysis is essential for monitoring sudden shifts in customer moods, detecting if complaints are on the rise, and for taking action before problems escalate. With sentiment analysis, you can monitor brand mentions on Twitter in real-time and gain valuable insights that tell you if you need to make updates.
2. **Scalability:** let's say you need to analyze hundreds of tweets mentioning your brand. While you could do that manually, it would take hours of manual processing, and as your data grows it would be impossible to scale. By performing Twitter sentiment analysis you can automate manual tasks and gain valuable insights in a very short time.
3. **Consistent Criteria:** analyzing sentiment in a text is subjective. When done manually. The same tweet may be viewed differently by two members of the same team. By training a machine learning model to perform sentiment analysis on Twitter data, you can use one set of criteria to analyze all your data, so results are consistent.

Integrated Academy of Management and Technology, Ghaziabad

SYSTEM SOFTWARE REQUIREMENT SPECIFICATION (SRS)

Below are the requirements used for running **TWITTER SENTIMENT ANALYSIS**

System Requirement

Download JetBrains PyCharm Community Edition 2019.2 x64

<https://www.jetbrains.com/pycharm/download/download-thanks.html?platform=windows&code=PCC>

Windows-Based Requirements

- Dual-core 64-bit processor
- 8 GB of memory
- Up to 24 GB of internal storage (PyCharm: 2.5GB+1GB for caches,)
- Windows 10, Windows 8.1 Update, Windows 8, and Windows 7.1

Library Requirements of PyCharm Platform

- Textblob(simple, pythonic text processing, Sentiment Analysis, part of speech tagging, noun phrase parsing)
- Tweepy(twitter library for Python)
- Matplotlib(Python plotting package) and MSVC-Runtime(Twitter Library)

Integrated Academy of Management and Technology, Ghaziabad

NH 24, Near Dasna Crossing, Adhyatmik Nagar, Udyog Kunj,
Ghaziabad, Uttar Pradesh 201009

Integrated Academy of Management and Technology, Ghaziabad

Code of Sentiment Analysis

#Importing Library/package

```
import matplotlib.pyplot as plt
import sys , tweepy
from textblob import TextBlob
```

```
def percentage(part, whole):
    return 100 * float(part) / float(whole)
```

#Entering Twitter API

```
consumerKey = "AAAA"
consumerSecret = "BBBB"
accessToken = "HHHH"
accessTokenSecret = "IIII"
```

#Authenticate API

```
auth = tweepy.OAuthHandler(consumer_key= consumerKey ,
consumer_secret=consumerSecret)
```

```
auth.set_access_token(accessToken,accessTokenSecret)
```

```
api = tweepy.API(auth)
```

#Fetch Tweets

```
searchTerm = input("Enter keyword hasgtab ")
noOfSearchTerm = int(input("Enter no of tweets" ))
```

```
tweets = tweepy.Cursor(api.search,q=
searchTerm).items(noOfSearchTerm)
```

Integrated Academy of Management and Technology, Ghaziabad

NH 24, Near Dasna Crossing, Adhyatmik Nagar, Udyog Kunj,
Ghaziabad, Uttar Pradesh 201009

Integrated Academy of Management and Technology, Ghaziabad

positive = 0
negative = 0
neutral = 0
polarity = 0

#Mining and Classified Tweets

for tweet in tweets:

```
    analysis = TextBlob(tweet.text)
    polarity += analysis.sentiment.polarity
    if analysis.sentiment.polarity == 0:
        neutral += 1
    elif analysis.sentiment.polarity < 0.00:
        negative += 1
    elif analysis.sentiment.polarity > 0.00:
        positive += 1
```

```
positive = percentage(positive, noOfSearchTerm)
negative = percentage(negative, noOfSearchTerm)
neutral = percentage(neutral, noOfSearchTerm)
positive = format(positive, '.2f')
negative = format(negative, '.2f')
neutral = format(neutral, '.2f')
```

```
if polarity == 0:
    print("Neutral")
elif polarity < 0:
    print("Negative")
elif polarity > 0:
    print("Positive")
```

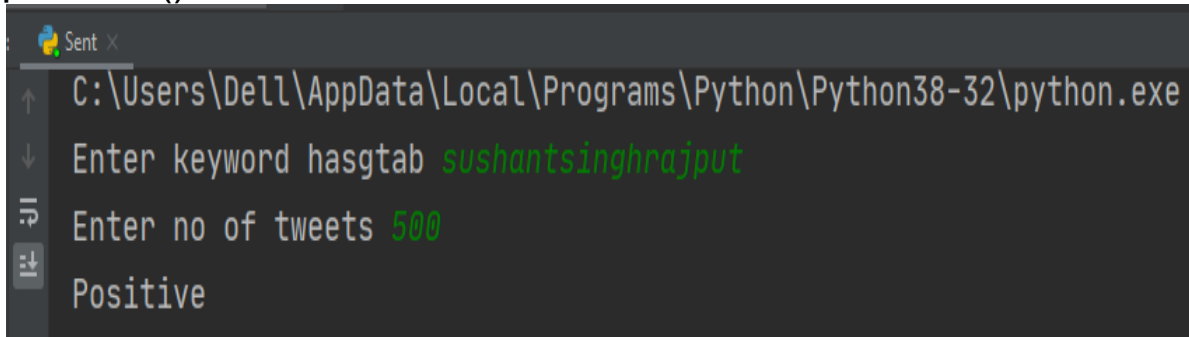
Integrated Academy of Management and Technology, Ghaziabad

NH 24, Near Dasna Crossing, Adhyatmik Nagar, Udyog Kunj,
Ghaziabad, Uttar Pradesh 201009

Integrated Academy of Management and Technology, Ghaziabad

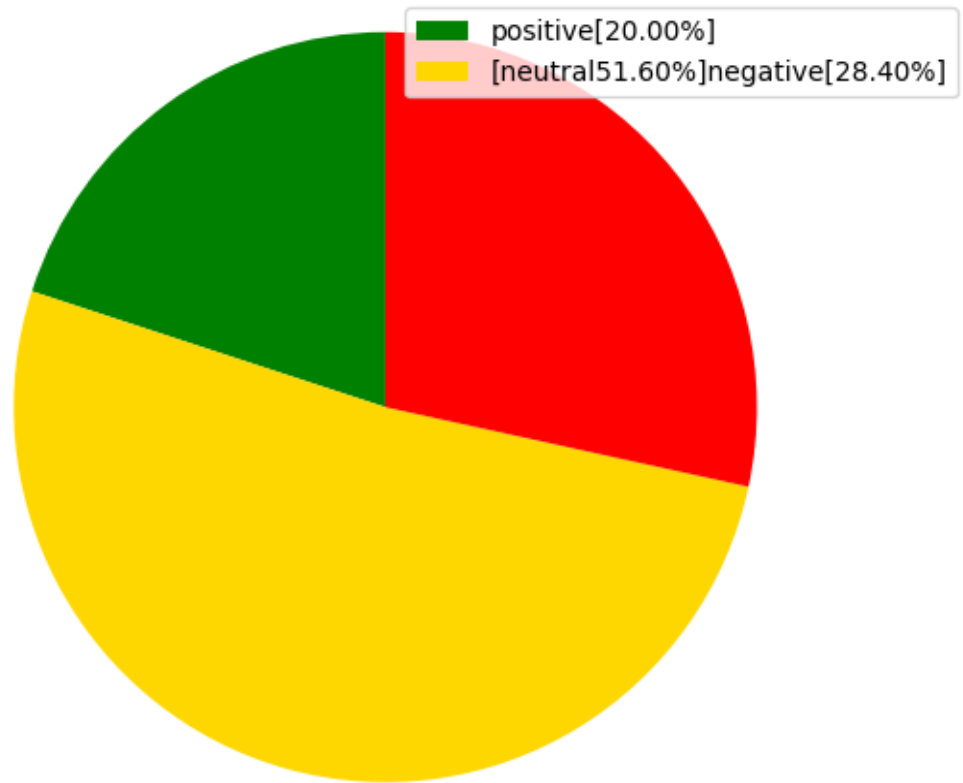
#Graphical Rep. of output

```
labels = ['positive[' + str(positive) + '%]', '[neutral' + str(neutral) + '%]' + 'negative[' + str(negative) + '%]']
sizes = [positive, neutral, negative]
colors = ['Green', 'gold', 'red']
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(noOfSearchTerm)+'Tweets.')
plt.axis('equal')
plt.tight_layout()
plt.show()
```



```
Sent x
C:\Users\De11\AppData\Local\Programs\Python\Python38-32\python.exe
Enter keyword hasgtab sushantsinghrajput
Enter no of tweets 500
Positive
```

How people are reacting on sushantsinghrajput by analyzing 500 Tweets.



Integrated Academy of Management and Technology, Ghaziabad
NH 24, Near Dasna Crossing, Adhyatmik Nagar, Udyog Kunj,
Ghaziabad, Uttar Pradesh 201009

Future Scope of the Project

Sentiment analysis is a uniquely powerful tool for businesses that are looking to measure attitudes, feelings and emotions regarding their brand. To date, the majority of sentiment analysis projects have been conducted almost exclusively by companies and brands through the use of social media data, survey responses and other hubs of user-generated content. By investigating and analyzing customer sentiments, these brands are able to get an inside look at consumer behaviors and, ultimately, better serve their audiences with the products, services and experiences they offer.

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. This forecast also predicts broader applications for sentiment analysis – brands will continue to leverage this tool, but so will individuals in the public eye, governments, nonprofits, education centers and many other organizations.

Deeper, Broader Insights from Sentiment Analysis

Sentiment analysis is getting better because social media is increasingly more emotive and expressive. A short while ago, Facebook introduced “Reactions,” which allows its users to not just ‘Like’ content, but attach an emoticon, whether it be a heart, a shocked face, angry face, etc. To the average social media user, this is a fun, seemingly silly feature that gives him or her a little more freedom with their responses. But, to anyone looking to leverage social media data for sentiment analysis, this provides an entirely new layer of data that wasn’t available before. Every time the major social media platforms update themselves and add more features, the data behind those interactions gets broader and deeper.

Greater Personalization for Audiences

As a result of deeper and better understanding of the feelings, emotions and sentiments of a brand or organization’s key, high-value audiences, members of these audiences will increasingly receive experiences and messages that are personalized and directly related to their wants and needs. Rather than segment markets based on age, gender, income and other surface demographics, organizations can further segment based on how their audience members actually *feel* about the brand or how they use social media. While some people shudder at the thought of companies learning more about them, more exact targeting

means that, in the near future, we will no longer be scratching our head wondering why we see advertisements for products we'd never dream of purchasing. In other words, the spray-and-pray advertising tactics are almost put to rest and there will be a time when every marketing message we see will be relevant and useful to us. Sentiment analysis is going to be a large contributing factor towards achieving this vision.

Not Just For Marketers and Brands

Again, sentiment analysis is on the verge of breaking into new areas of application. While we will likely always think of it first in terms of the traditional marketing sense, the world has already seen a few ways that sentiment analysis can be used in other areas. Social media analytics helped predict and explain the emotions of concerned parties behind Brexit and the 2016 US election, which has spurred a number of non-brand organizations to investigate how sentiment analysis can be used to predict outcomes and map out the emotional landscape of people, voters and the like. Additionally, businesses are looking at ways that sentiment analysis can be used outside of their marketing and PR departments. Sentiment analysis simply looks more popular in the future.

Algorithm-Based Sentiment Analysis Plateaus

Algorithms have long been at the foundation of most forms of analytics, including social media and sentiment analysis. With recent years bringing big leaps in machine learning and artificial intelligence, many analytics solutions are looking to these technologies to replace algorithms. Unfortunately for organizations looking to leverage sentiment analysis to measure

audience emotions, machine learning isn't yet ready to tackle the complex nuances of text and how we talk, especially on social media channels that are rife with slang, sarcasm, double meanings and misspellings. These make it difficult for artificial intelligence systems to accurately sort and classify sentiments on social media. And, with any analysis project, accuracy is crucial. It is uncertain if machine learning will progress to the point that it *is* capable of accurately analyzing text, or if sentiment analysis projects will have to find a new basis to avoid the current plateau of algorithms. Some social media analytics solutions have begun taking a more human approach to deciphering the often ambiguous nature of text, but this can be time consuming.

Integrated Academy of Management and Technology, Ghaziabad

NH 24, Near Dasna Crossing, Adhyatmik Nagar, Udyog Kunj,
Ghaziabad, Uttar Pradesh 201009