MADHAV INSTITUTE OF TECHNOLOGY & SCIENCE, GWALIOR
(A Govt. Aided UGC Autonomous & NAAC Accredited Institute Affiliated to RGPV, Bhopal)

Project Report
on
SENTIMENT ANALYSIS ON TWITTER DATA

A research paper submitted in partial fulfilment of the requirement for the degree of

BACHELOR OF TECHNOLOGY
in
INFORMATION TECHNOLOGY

Submitted by:

Ankita Gupta

0901IT181009

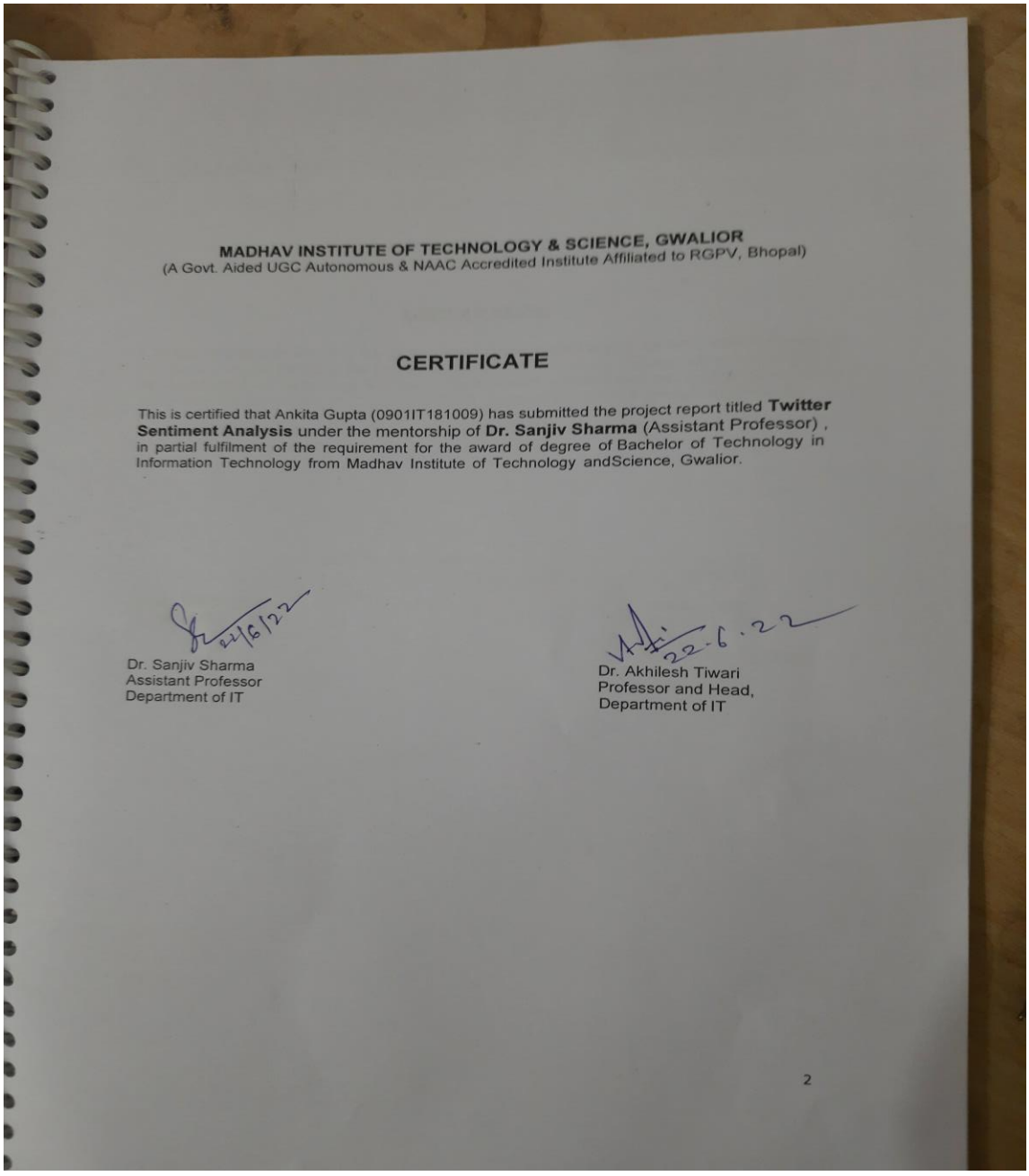
Project – 160801

Faculty Mentor
Dr. Sanjiv Sharma
Assistant Professor

Submitted to:

DEPARTMENT OF INFORMATION TECHNOLOGY
MADHAV INSTITUTE OF TECHNOLOGY & SCIENCE
GWALIOR – 474005

MAY-JUNE 2022



MADHAV INSTITUTE OF TECHNOLOGY & SCIENCE, GWALIOR
(A Govt. Aided UGC Autonomous & NAAC Accredited Institute Affiliated to RGPV, Bhopal)

CERTIFICATE

This is certified that Ankita Gupta (0901IT181009) has submitted the project report titled **Twitter Sentiment Analysis** under the mentorship of **Dr. Sanjiv Sharma** (Assistant Professor) , in partial fulfilment of the requirement for the award of degree of Bachelor of Technology in Information Technology from Madhav Institute of Technology andScience, Gwalior.

24/6/22

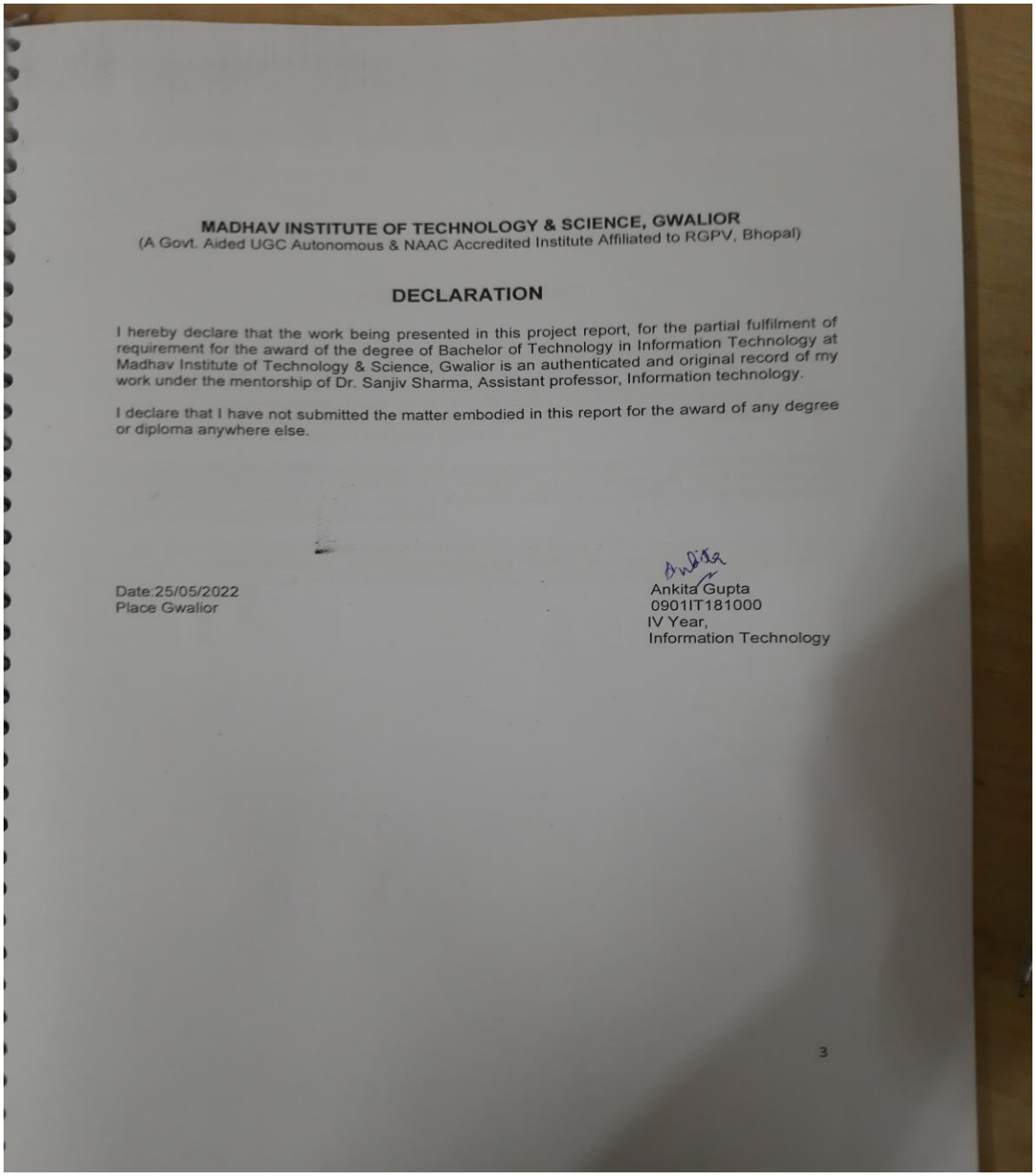
Dr. Sanjiv Sharma
Assistant Professor
Department of IT

22.6.22

Dr. Akhilesh Tiwari
Professor and Head,
Department of IT

2

MADHAV INSTITUTE OF TECHNOLOGY & SCIENCE, GWALIOR
(A Govt. Aided UGC Autonomous & NAAC Accredited Institute Affiliated to RGPV, Bhopal)



MADHAV INSTITUTE OF TECHNOLOGY & SCIENCE, GWALIOR
(A Govt. Aided UGC Autonomous & NAAC Accredited Institute Affiliated to RGPV, Bhopal)

DECLARATION

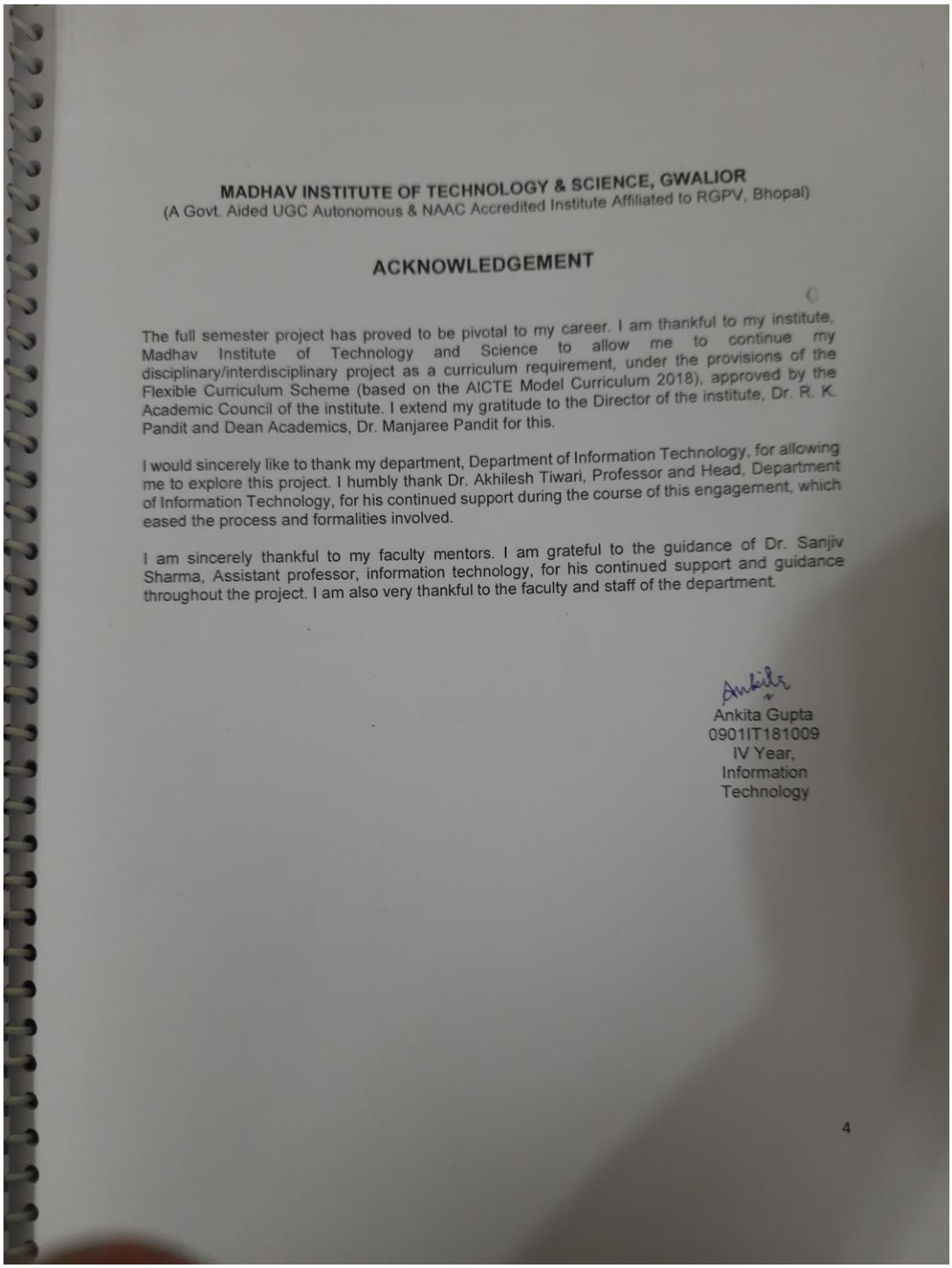
I hereby declare that the work being presented in this project report, for the partial fulfilment of requirement for the award of the degree of Bachelor of Technology in Information Technology at Madhav Institute of Technology & Science, Gwalior is an authenticated and original record of my work under the mentorship of Dr. Sanjiv Sharma, Assistant professor, Information technology.

I declare that I have not submitted the matter embodied in this report for the award of any degree or diploma anywhere else.

Date:25/05/2022
Place Gwalior

Ankita Gupta
0901IT181000
IV Year,
Information Technology

3



MADHAV INSTITUTE OF TECHNOLOGY & SCIENCE, GWALIOR
(A Govt. Aided UGC Autonomous & NAAC Accredited Institute Affiliated to RGPV, Bhopal)

ACKNOWLEDGEMENT

The full semester project has proved to be pivotal to my career. I am thankful to my institute, Madhav Institute of Technology and Science to allow me to continue my disciplinary/interdisciplinary project as a curriculum requirement, under the provisions of the Flexible Curriculum Scheme (based on the AICTE Model Curriculum 2018), approved by the Academic Council of the institute. I extend my gratitude to the Director of the institute, Dr. R. K. Pandit and Dean Academics, Dr. Manjaree Pandit for this.

I would sincerely like to thank my department, Department of Information Technology, for allowing me to explore this project. I humbly thank Dr. Akhilesh Tiwari, Professor and Head, Department of Information Technology, for his continued support during the course of this engagement, which eased the process and formalities involved.

I am sincerely thankful to my faculty mentors. I am grateful to the guidance of Dr. Sanjiv Sharma, Assistant professor, information technology, for his continued support and guidance throughout the project. I am also very thankful to the faculty and staff of the department.

Ankita Gupta
0901IT181009
IV Year,
Information
Technology

TABLE OF CONTENTS

Abstract:

Twitter is one the most popular social media platform for micro-blogging, where people express their opinion freely and un-conservatively about anything and everything. There is a tremendous amount of pre-existing data on the web for researchers, and a vast amount of data is generated daily. There are about 500 Million new tweets on the platform, making it very diverse. Analyzing this data continues to be a popular and fruitful field of research in the recent past. The Twitter data is highly unstructured and has high entropy, so this paper comparatively analyzes various machine learning techniques, which makes it easy to understand and analyze the opinion of people, also known as sentiment analysis/opinion mining where the tweets(Twitter messages) are classified as positive or negative. We have also discussed difficulties that are faced in general while doing sentiment analysis on Twitter data.

KEYWORDS:

Sentiment Analysis, Machine Learning, Support vector Machine(SVM), Lexicons, VADER, Logistic Regression, Neural Networks, Decision trees, Gradient Boosting

सार:

ट्विटर माइक्रो-ब्लॉगिंग के लिए सबसे लोकप्रिय सोशल मीडिया प्लेटफार्मस में से एक है, जहां लोग किसी भी चीज और हर चीज के बारे में अपनी राय खुलकर और गैर-रूढ़िवादी रूप से व्यक्त करते हैं। शोधकर्ताओं के लिए वेब पर पहले से मौजूद डेटा भारी मात्रा में उपलब्ध है, और प्रतिदिन बड़ी मात्रा में डेटा उत्पन्न होता है। प्लेटफ़ॉर्म पर लगभग 50 करोड़ नए ट्वीट हैं, जो इसे बहुत विविध बनाते हैं। इस डेटा का विश्लेषण हाल के दिनों में अनुसंधान का एक लोकप्रिय और उपयोगी क्षेत्र बना हुआ है। ट्विटर डेटा अत्यधिक असंरचित है और इसमें उच्च एन्ट्रॉपी है, इसलिए यह शोषपत्र तुलनात्मक रूप से विभिन्न मशीन लर्निंग की तकनीकों का विश्लेषण करता है, जिससे लोगों की राय को समझना और विश्लेषण करना आसान हो जाता है, जिसे भावना विश्लेषण / राय खनन के रूप में भी जाना जाता है जहां ट्वीट्स (ट्विटर संदेश) हैं सकारात्मक या नकारात्मक के रूप में वर्गीकृत। हमने उन कठिनाइयों पर भी चर्चा की है जो आम तौर पर ट्विटर डेटा पर भावना विश्लेषण करते समय आती हैं।

LIST OF FIGURES:

LIST OF TABLES

CHAPTER 1 - Introduction:

With the widespread adoption of the internet, people can express their views, opinions, and how a person is affected by another's thoughts. These are predominantly done through social media websites, blogs, online forums, etc. Currently, billions of people use social media sites like Facebook, Twitter, Instagram, and others regularly. People share numerous things that go on in their lives and affect them through different mediums like images, gifs, text, emojis, etc.

These platforms constantly generate a large volume of sentiment-rich data in tweets, micro-blogging, blogs, comments, reviews of products, etc. Tweets depict an individual's opinion on various socio-economic events and his/her personal beliefs, thoughts, and outlook toward the world.

The amount of content generated by Twitter on a daily basis is too vast for a normal user to analyze. For this purpose, various sentiment analysis techniques are widely used, some of which are discussed here.

Sentiment analysis (SA) is an integral part of Natural Language Processing. Natural Language Processing is a technique that makes our machines understand human language in various forms, whether it is spoken words or written texts. Sentiment analysis, to be exact, tells whether the data in any form is positive or negative. Various parts of data affect and do not affect this classification in many ways, which we will see further in the paper.

Here we discuss different steps in obtaining the sentiment of data and various techniques to do the same.

Various steps include collecting the tweets, analyzing the data using various methods, structuring the data, cleaning or preprocessing the data, testing different models on the training data, and at last, using precision measures to analyze and compare the results obtained.

We have obtained close to 90 percent accuracy using numerous machine learning models like SVM, Logistic regression, and many more.

CHAPTER 2 - Literature Review:

In last decade a lot of work has been done in the field of “Sentiment Analysis on Twitter“ by multiple researchers. Earlier it was binary classification where the text is classifies as either “positive” or “negative”. There are various text mining techniques used to mine the data.

The authors [1] studied the government of India policy of Demonetization. They performed sentiment analysis on tweets on the basis of both nation-wide and state-wise impact. They collected the corpus of tweets using Twitter API using particular hashtags. Research based on geo-location (State-wise tweets are gathered). The sentiment analysis is performed to categorize the tweet into five categories - happy, sad, very sad, very happy, neutral, and no data.

The authors [2] performed sentiment analysis on the Amazon product reviews dataset on both sentence and review levels. They used Naïve Byes, SVM, and Random Forest to generate the polarity of reviews, where Random Forest generated the best results with 92 percent accuracy, with 90 percent for Naïve Byes and 82 percent for the SVM classifier.

The author [3] Performed sentiment analysis on Twitter data and classified the tweets into three categories - Positive, Negative, and Neutral. They pre-processed the dataset and then applied various machine learning models like Naive Byes, SVM, and Maximum entropy. They gained maximum accuracy using the Naive Bayes Algorithm, which measures 89.5%.

The authors [4] conducted sentiment analysis on tweet corpus about the electronic products and classified the data into three lables - Positive, Negative, and Neutral. The authors used a machine learning-based approach and used Naive Bayes, SVM, Maximum Entropy, and Ensamble Learning approach where all classifiers gave almost the same precision.

The authors [5] performed sentiment analysis on tweet corpus of movie reviews based on both Hollywood and Bollywood movies. They classified the tweet using machine learning approach where they used Naïve Bayes and SVM classifier and gained accuracy of 90 percent using SVM and 88 percent using Naïve Bayes classifier.

The Following table shows the works of various authors on Sentiment Analysis having Twitter dataset.

Author & Year	Dataset	Techniques	Accuracy
Divakr (2014) [3]	Tweets about customer reviews	NB Maximum Entropy SVM Word net	88.2% 83.8% 85.5% 89.9%
Neethu M. S. (2013) [4]	Tweets about electronic products	Navive Bayes SVM Maximum Entropy Essembled	89.5% 90% 90% 90%
Justin Zhan(2015)[2]	Amazon product reviews	Random Forest NB SVM	92% 90% 82%
Akshay Amolik (2015)[5]	Twitter posts About movie reviews	Naïve Bayes SVM	88% 90%

Table 1: Related work

CHAPTER 3 - Proposed work

Flow chart-

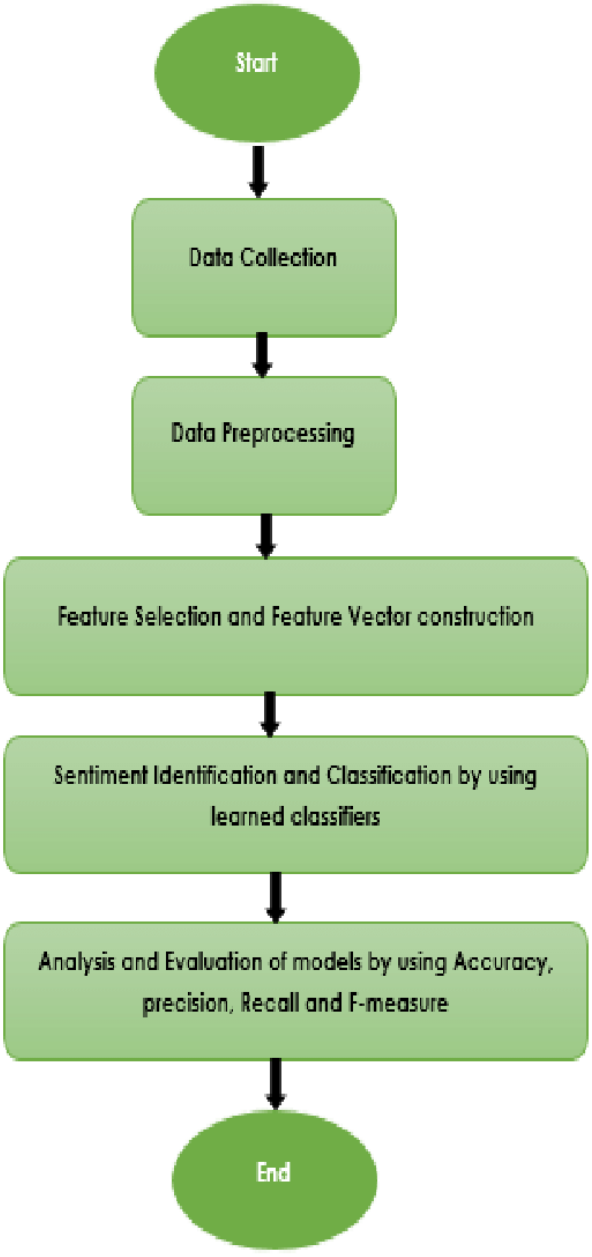


Fig 1: Flow Chart of model

SENTIMENT ANALYSIS

Sentiment analysis is the procedure that mines the opinion polarity of any document or text as a part of the Natural Language Processing (NLP) technique. The sentiments can be classified into many categories, for example, positive, happy, neutral, sad, etc., depending upon the use case and the problem we are trying to analyze.[6]

Sentiment analysis can be done in the following steps:

- Data loading and visualization
- Preprocessing of data
- Text vectorization or representation
- Developing Model

4.1 DATA LOADING AND VISUALIZATION

The dataset used of around 18000 tweets which have labeled sentiments along with tweet text and tweet id which is sourced from Kaggle in the form of CSV file. The Data includes 9897 positive and 8830 negative tweets.

	textID	tweet_text	sentiment
4414	1962463307	@RobertsEmma You don't reply to none of your f...	negative
5615	1963383772	Oh why do I keep stumbling about deathfic WITH...	negative
2759	1961049714	@Cheapflights_uk o2 - train from Bristol to L...	negative
1524	1957729651	my knee hurts!	negative

Figure 2 – view of dataset used.

The concept of using pictures is to comprehend data that has been used throughout history. Data visualization is the graphic representation of data in a pictorial or graphical format(Eg. charts, graphs,etc). Data visualization tools deliver a convenient way to see and understand trends, patterns in data, and outliers.[7]

Here we have visualized our data in 2 ways:

1. In the form of a pie chart using Matplotlib.
2. In the form of word clouds.

Matplotlib.pyplot helps understanding and representing how tweets are distributed over the dataset

Word Cloud visualize the individual words of data based on the frequency of occurring. Bigger the font size of word depicts higher occurrence.

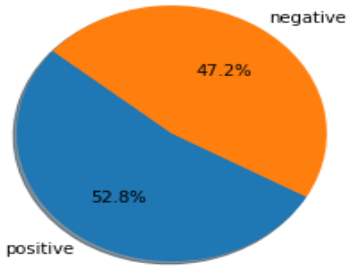


Figure 3 - Pie chart of our used dataset created using Matplotlib library in python.

Figure 4 - word cloud built on positive tweets based of dataset.

4.2 PREPROCESSING OF DATA

People from different geographic/social /economic/political backgrounds use Twitter. The raw data having polarity is exposed to inconsistent writing and redundant information(s) for sentiment analysis.. So, the following preprocessing was done before training the models.

- All URLs were removed (e.g., www.facebook.com),
- All hashtags were removed
- All user names were removed (e.g., @Thorin)
- Remove the RT tags, which refer to re-tweets in the dataset.
- The sequence of repeated characters and punctuations were removed. (e.g., looooooooook -> look, hey!!!!!!!! -> hey!)
- Lowercasing each letter of data.
- Replaced all the emoticons with text.

- Removed all punctuations, symbols, and numbers.
- Word contraction. (e.g., I'm -> I am)
- Remove Stop Words(e.g., is, am, are)
- **Tokenization**: Separating strings of a tweet/sentence into smaller parts, mostly individual words. We used NLTK's 'word_tokenize' module and also tried our custom tokenizer
- **Stemming** - Stemming is the process of converting a word to its root form(learned -> learn).
- **Lemmatization** - lemmatization looks beyond word reduction and considers a language's full vocabulary to apply a morphological analysis to words. E.g., The lemma of 'was' is 'be,' and the lemma of 'mice' is 'mouse.'[8]

4.3 TEXT VECTORIZATION

Text Vectorization is the process of converting text into vectors, so we can use them as an input to our models The methods we used are-

- **Positive/Negative frequency**
- **Bag of words or Count Vectorizer**

A bag of words representation of text counts the occurrence of words within a document.[9]

```
array([[0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0],
       [0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1],
       [0, 0, 0, 3, 0, 0, 1, 0, 0, 0, 0, 0, 0],
       [0, 0, 1, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0],
       [0, 1, 0, 0, 1, 1, 0, 0, 1, 0, 0, 0, 0],
       [1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 1, 0]])
```

Figure 5 – Example of Bag of word matrix on tweet corpus of size 6.

- **TF-IDF(Term Frequency – Inverse Document Frequency)**

TF-IDF(Term Frequency – Inverse Document Frequency). The term frequency is the number of occurrences of a specific word in a document. Term frequency indicates how vital a particular term is in a document. Since the no of occurrence of a particular word by itself does not add much information, without the knowledge of the distribution across all the documents, using tf-idf helps considerably in most cases

Inverse document frequency (IDF) is the weight of a term. It seeks to lower a term's weight if it is dispersed throughout all the documents.

idf(t) = log [n / (df(t) + 1)]

TF-IDF was the best performing in almost every machine learning models used including SVM, Logistic Regression, Naïve Byes and Decision Trees algorithms.

```
array([[0.        , 0.        , 0.        , 0.        , 0.        ,
        0.        , 0.56921261, 0.        , 0.        , 0.82219037,
        0.        , 0.        , 0.        ],
       [0.        , 0.        , 0.        , 0.        , 0.        ,
        0.        , 0.        , 0.70710678, 0.        , 0.        ,
        0.        , 0.        , 0.70710678],
       [0.        , 0.        , 0.        , 0.96260755, 0.        ,
        0.        , 0.27089981, 0.        , 0.        , 0.        ,
        0.        , 0.        , 0.        ],
       [0.        , 0.        , 0.68172171, 0.55902156, 0.        ,
        0.        , 0.47196441, 0.        , 0.        , 0.        ,
        0.        , 0.        , 0.        ],
       [0.        , 0.52182349, 0.        , 0.        , 0.42790272,
        0.52182349, 0.        , 0.        , 0.52182349, 0.        ,
        0.        , 0.        , 0.        ],
       [0.52182349, 0.        , 0.        , 0.        , 0.42790272,
        0.        , 0.        , 0.        , 0.        , 0.        ,
        0.52182349, 0.52182349, 0.        ]])
```

Figure 6 – TF-IDF matrix on tweet corpus of size 6

4.5 DEVELOPMENT OF MODEL

All models were made with a test/train split ratio of 0.2 and with random state=0 to keep experimental conditions the same to compare results. Any changes in random state/ratio may/may not induce tiny changes in model performance.

All the following metrics mentioned henceforth are for the test set.

We calculate accuracy by dividing the number of correct predictions (the corresponding diagonal in the matrix) by the total number of samples.

$$\text{Accuracy} = \frac{(TP + TN)}{(TP + FP + TN + FN)}$$

Figure – 7 formula of accuracy

4.5.1 LEXICON BASED APPROACH

The first approach that can be used for Sentiment Analysis is the lexicon-based technique. Here, each term is labeled with sentiment score which helps in determining the polarity of tweet.

We have achieved an accuracy of 67.77% while using TF-IDF as vectorization method, which was found to be the best. We will take this as a base model for further speculations.

4.5.2 DICISION TREES

The tree consists of root, internal, and leaf nodes. The tree is then used in categorizing unknown data points. Since we can make deeper/wider trees that completely classify the given training dataset, the trees are prone to overfitting, and we must carefully decide the size/pruning process. The decision tree classification approach is executed in two stages: tree building and tree pruning

This algorithm generated the best results in prediction of the polarity of tweets with the TF-IDF vectorization technique with the accuracy of around 68% while it was tested on positive/negative frequency and Bag of words too but didn't generate satisfactory results. We have obtained an increase of 2% from our base model.

4.5.3 SVM (SUPPORT VECTOR MACHINE)

SVMs try to classify a given datapoint by finding the relationships between features in a higher dimension hyper-plane to find the relations that distinguish the different classes.

The model finds those hyperplane(s) by transforming the input data using mathematical transformations known as 'Kernels.' Commonly used kernels are of 2 types

1) Linear
2) Non-Linear (polynomial, rbf, sigmoid)

We can also define our custom kernels if required, although the radial basis function (rbf) works well for most non-linear problems by varying its hyper-parameters. In our problem, linear decision boundary was working as well as others so we have used linear SVM here.

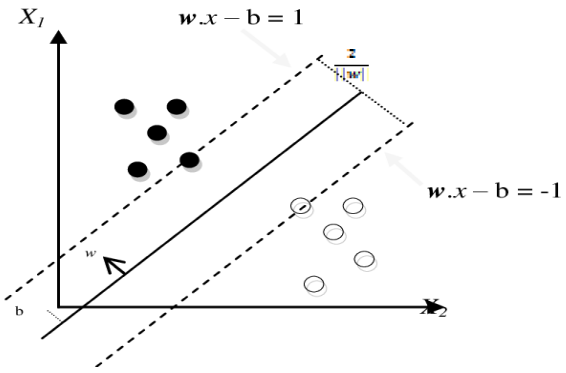


Figure 8 - Visual Representation of optimal hyperplane between two classes, and it's supporting vectors.

This algorithm generated the best results in prediction of the polarity of tweets with the TF-IDF vectorization technique with the accuracy of around 88.2%. We have obtained an increase of around 21% from our base model.

4.5.4 KNN (K- NEAREST NEIGHBOURS)

KNN is an algorithm that categorizes the given datapoint by interpolating the labels from the K nearest matches in training data. The weights of each neighbor may be linear or some function of distance.

This algorithm generated the best results in prediction of the polarity of tweets with the Bag of words technique with the accuracy of around 79.95%, and 79.87% using Positive/Negative frequency approach. We have obtained an increase of around 13% from our base model.

4.5.5 LOGISTIC REGRESSION

Logistic regression: Logistic regression is a widely used approach to solve binary classification problems. Sentiment Analysis is one of many fields where logistic regression may be applied. The idea behind the method is to calculate the probability by fitting the logit function to the given data.

$$p(x) = \frac{1}{1 + e^{-(x-\mu)/s}}$$

This method generated the best results in prediction of the polarity of tweets with the TF-IDF vectorization technique with the accuracy of around 88.57%. We have obtained an increase of around 21% from our base model.

4.5.6 GAUSSIAN NAÏVE BAYES

The Naive Bayes algorithm helps us find the probability of the tweet having a particular sentiment by fitting the feature values to a probability distribution and multiplying the probabilities for those features for a given input. Here we fitted the feature values to a Gaussian distribution. The Naive Bayes classifier has an inherent simplicity in its implementation and interpretation and is generally fast. For some problems, it gets the job done. However, the assumption of independent occurrence of features may cause a problem in cases where the features are highly dependent.

This method generated the best results in predicting the polarity of tweets with the TF-IDF vectorization technique with an accuracy of around 71%. We have obtained an increase of about 4% from our base model.

4.5.7 ENSAMBLE APPROACH (SVM, LOGISTIC REGRESSION, NAÏVE BAYES)

Here, we have built a voting classifier using Sklearn library and used SVM, Logistic Regression, Gaussian Naïve Bayes for predictions of polarity of tweet corpus. We have achieved an accuracy of 88% while using TF-IDF as the vactorization method.

MACHINE LEARNING APPROACH	GAINED ACCURACY
DICISION TREES	**68%**
SVM	**88.2%**
KNN	**79.95%**
Logistic Regresion	**88.57%**
Gaussian Naïve Byes	**71%**
SVM+Logistic Regression+Naïve Byes	**88%**

Table 2 – Summary of model's accuracy

Chapter 4 – snapshots of GUI for project

We have used Streamlit library for development of Graphical User Interface(GUI).

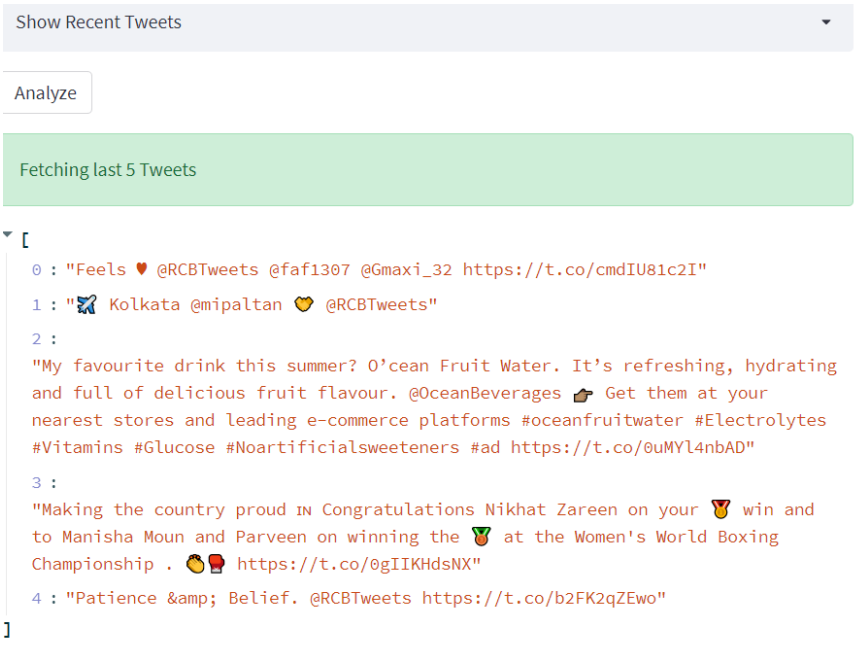


Figure 9 – feature for generating recent tweets

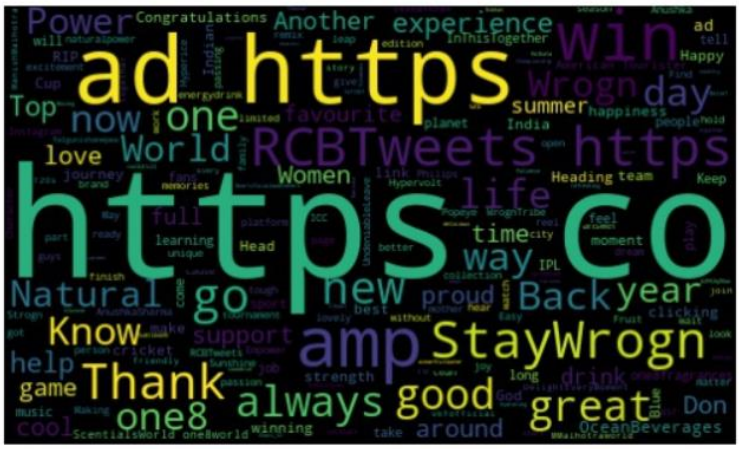


Figure 10 – Generation of word cloud using recent tweets

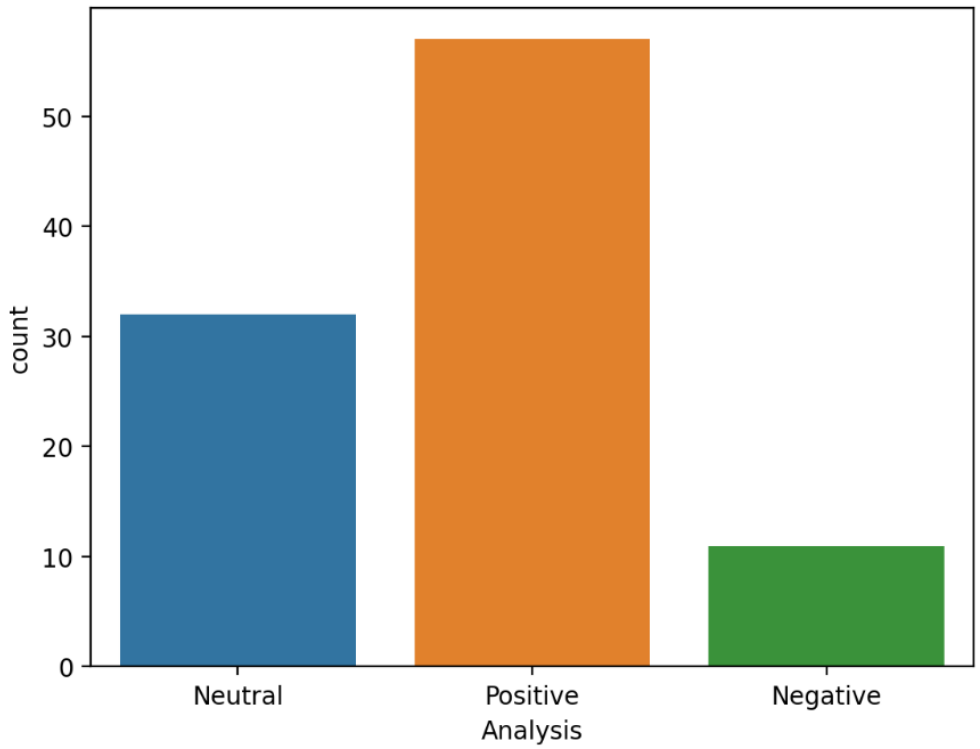


Figure 11 – graph showing tweets polarity

Fetching Last 100 Tweets

	Tweets	Subjectivity	Polarity	Analysis
0	Feels ♥ 1307 _32	0.0000	0.0000	Neutral
1	✈ Kolkata 🤝	0.0000	0.0000	Neutral
2	My favourite drink this summe…	0.8500	0.6167	Positive
3	Making the country proud IN C…	0.7167	0.7000	Positive
4	Patience & Belief.	0.0000	0.0000	Neutral
5	Cleans day. 60 and 70 kg. Easy-…	0.0000	0.0000	Neutral
6	With natural power comes nat…	0.6100	0.2500	Positive
7	"Results aren't in your hands, …	0.0000	0.0000	Neutral
8	Livin' that PUMA life 🐆⬛ Sh…	0.0000	0.0000	Neutral
9	A historic achievement and a …	0.5833	0.1667	Positive

Figure 12 – sentiments in tabular form

Conclusion:

In this project, we have built a simple GUI demonstrating the implementation of sentiment analysis on twitter in real time. We have achieved it using various Machine Learning models and a Lexicon based model, where we achieved 88.57% accuracy on the testing dataset.

References:

[1] Prabhsimran, Singh, Ravinder Singh Sawhney, and Karanjeet Singh Kahlon. 2017. Sentiment analysis of demonetization of 500 & 1000 rupee banknotes by Indian government

[2] Fang, and Justin Zhan, Xing. 2015. "Sentiment analysis using product review data."

[3] Gautam, Geetika, and Divakar Yadav. "Sentiment analysis of twitter data using machine learning approaches and semantic analysis." Contemporary computing (IC3), 2014.

[4] Neethu, M. S., and R. Rajasree. "Sentiment analysis in twitter using machine learning techniques." Computing, Communications and Networking Technologies (ICCCNT), 2013 Fourth International Conference on. IEEE, 2013.

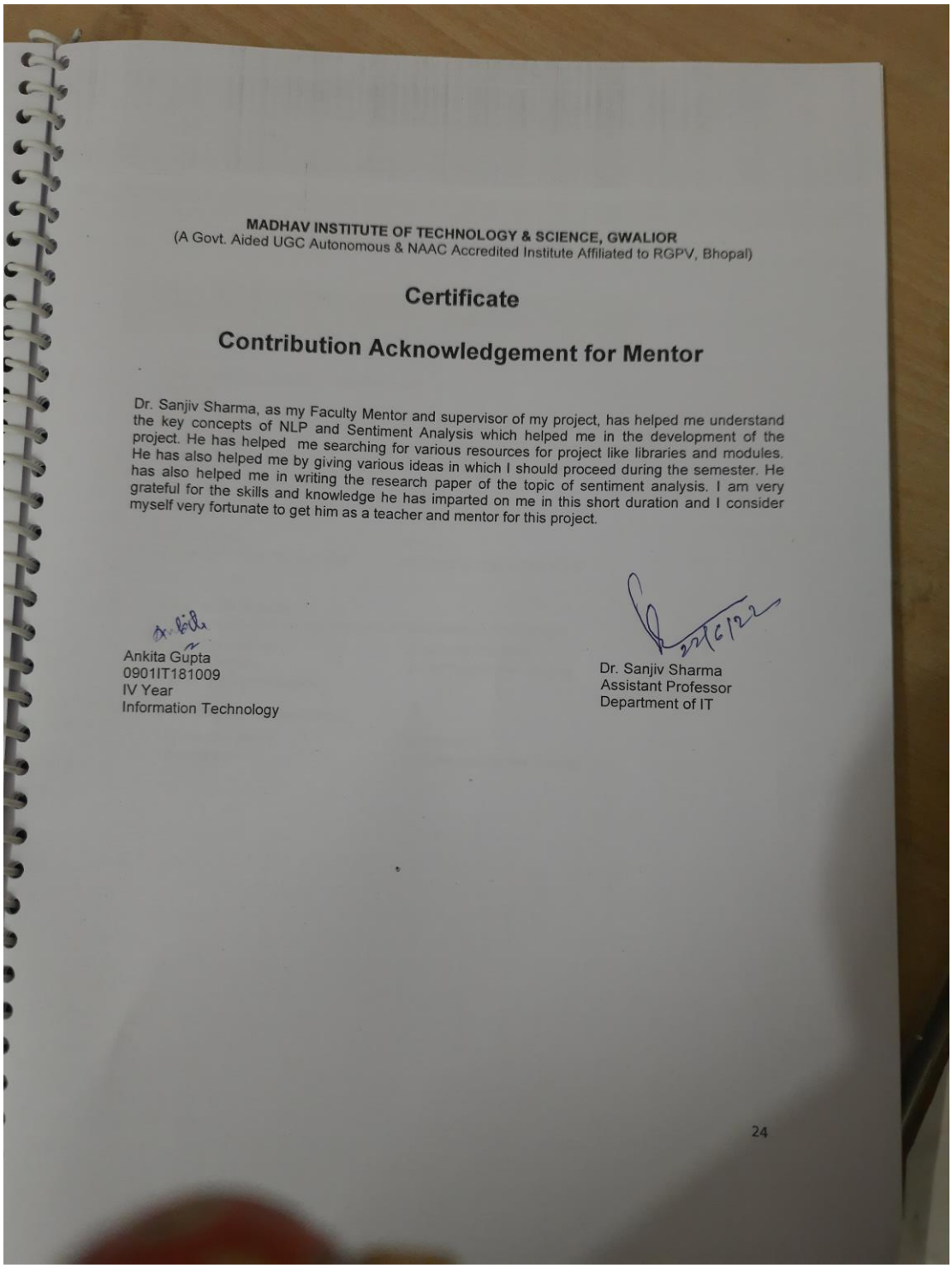
[5] Amolik, Akshay, et al. "Twitter sentiment analysis of movie reviews using machine learning techniques." International Journal of Engineering and Technology 7.6 (2016): 1-7.

[6]https://monkeylearn.com/sentiment-analysis/

[7]https://www.geeksforgeeks.org/what-is-data-visualization-and-why-is-it-important/#:~:text=Data%20visualization%20is%20the%20graphical,patterns%20in%20data%2C%20and%20outliers.

[8] https://towardsdatascience.com/stemming-vs-lemmatization-2daddabcb221#:~:text=Stemming%20and%20Lemmatization%20both%20generate,words%20which%20makes%20it%20faster.

MADHAV INSTITUTE OF TECHNOLOGY & SCIENCE, GWALIOR
(A Govt. Aided UGC Autonomous & NAAC Accredited Institute Affiliated to RGPV, Bhopal)



MADHAV INSTITUTE OF TECHNOLOGY & SCIENCE, GWALIOR
(A Govt. Aided UGC Autonomous & NAAC Accredited Institute Affiliated to RGPV, Bhopal)

Certificate

Contribution Acknowledgement for Mentor

Dr. Sanjiv Sharma, as my Faculty Mentor and supervisor of my project, has helped me understand the key concepts of NLP and Sentiment Analysis which helped me in the development of the project. He has helped me searching for various resources for project like libraries and modules. He has also helped me by giving various ideas in which I should proceed during the semester. He has also helped me in writing the research paper of the topic of sentiment analysis. I am very grateful for the skills and knowledge he has imparted on me in this short duration and I consider myself very fortunate to get him as a teacher and mentor for this project.

Ankita Gupta
0901IT181009
IV Year
Information Technology

Dr. Sanjiv Sharma
Assistant Professor
Department of IT

24

turnitin

Similarity Report ID: oid:28506:19226384

PAPER NAME	AUTHOR
report_Sentiment_analysis_using_twitter_data.pdf	a g

WORD COUNT	CHARACTER COUNT
3378 Words	18297 Characters
PAGE COUNT	**FILE SIZE**
24 Pages	1.2MB
SUBMISSION DATE	**REPORT DATE**
Jun 22, 2022 12:31 PM GMT+5:30	Jun 22, 2022 12:32 PM GMT+5:30

9% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

- 8% Internet database
- 0% Publications database
- Crossref database
- Crossref Posted Content database
- 7% Submitted Works database

Excluded from Similarity Report

- Bibliographic material
- Quoted material
- Cited material
- Small Matches (Less then 14 words)

turnitin

Similarity Report ID: oid:28506:19226384

9% Overall Similarity

Top sources found in the following databases:

- 8% Internet database
- Crossref database
- 7% Submitted Works database
- 0% Publications database
- Crossref Posted Content database

TOP SOURCES

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	**Madhav Institute of Technology & Science on 2019-05-14** Submitted works	2%
2	**coursehero.com** Internet	2%
3	**github.com** Internet	1%
4	**mitsgwalior.in** Internet	1%
5	**Madhav Institute of Technology & Science on 2019-05-02** Submitted works	<1%
6	**Queen Mary and Westfield College on 2022-05-20** Submitted works	<1%
7	**hre.iitr.ac.in** Internet	<1%
8	**rucore.libraries.rutgers.edu** Internet	<1%