A Project Report on

WOMEN SAFETY SENTIMENT ANALYSIS
BASED ON TWITTER DATA USING ML.

Submitted to

DR. APJ ABDUL KALAM TECHNICAL UNIVERSITY
LUCKNOW

For the Partial Fulfillment of the Requirements for the Degree of

BACHELOR OF TECHNOLOGY IN COMPUTER SCIENCE
AND ENGINEERING

Submitted to:-

Department of Computer Science and Engineering

Shanti Institute of Technology , Baghpat Road (Meerut)

Mr. Naveen Kumar (HOD CSE)

Submitted by:-

- Ayush kumar(1903700100003)   - Km Minakshi(2003700109003)

DECLARATION

We hereby affirm that the project in this report titled "WOMEN SAFETY SENTIMENT ANALYSIS BASED ON TWITTER DATA USING ML", is an authentic and original project conducted by us at the Department of Computer Science & Engineering, Shanti Institute Of technology Kurali Meerut. This work has been submitted to fulfill the requirements for the Bachelor of Technology degree in Computer Science & Engineering awarded by Dr. A.P.J. Abdul Kalam Technical University in Lucknow. The contents of this report are true to the best of our knowledge, and we confirm that this project has not been previously submitted for the purpose of obtaining any other degree or graduation. Signature: AYUSH KUMAR Roll No-1903700100003 and KM MINAKSHI Roll No – 2003700109003

Signature:-

AYUSH KUMAR (1903700100003)

KM MINAKSHI (2003700109003)

<u>CERTIFICATE</u>

This is to certify the project report entitled "WOMEN SAFETY SENTIMENT ANALYSIS BASED ON TWITTER DATA USING ML" Done by Ayush kumar (1903700100003) and Km Minakshi (2003700109003) is an original Work carried out by them in department of computer science and engineering, Shanti Institute of technology Kurali Meerut under my guidance. The matter Embodied in this project work has not been submitted earlier for the award of any Degree or diploma to the best of my knowledge or belief.

DATE:

MR. NAVEEN KUMAR   MR. NITIN SIR

Signature of the Supervisor   Head of Department

<u>ABSTRACT</u>

Women and girls have been experiencing a lot of violence and harassment in public places in various cities starting from stalking and leading to abuse harassment or abuse assault. This project basically focuses on the role of social media in promoting the safety of women in Indian cities with special reference to the role of social media websites and applications including Twitter plat form Facebook and Instagram. This project also focuses on how a sense of responsibility on part of Indian society can be developed the common Indian people so that we should focus on the safety of women surrounding them. Tweets on Twitter which usually contains images and text and also written messages and quotes which focus on the safety of women in Indian cities can be used to read a message amongst the Indian Youth Culture and educate people to take strict action and punish those who harass the women. Twitter and other Twitter handles which include hash tag messages that are widely spread across the whole globe sir as a platform for women to express their views about how they feel while we go out for work or travel in a public transport and what is the state of their mind when they are surrounded by unknown men and whether these women feel safe or not? . This may assist our studies to overcome the emotions of humans around us.

ACKNOWLEDGEMENTS

We express our heartfelt gratitude to the divine for guiding us throughout this project, enabling us to bring it to a successful conclusion. With humble hearts, we pray for the continuous guidance and blessings of the almighty in all our endeavors. Our deepest appreciation to our project guide Mr. Naveen kumar, whose invaluable guidance and wisdom have been instrumental in navigating the challenges we encountered during the project. His extensive knowledge and support have been of light throughout the project's duration. We would also like to extend our sincere thanks to Mr. Nitin sir, head of department, for a being a constant source of support and assistance in very possible way during the project. His presence and guidance have been invaluable in our journey. We are immensely grateful to all the individuals who have directly or indirectly supported us during the completion of this project. Their contribution and encouragement have been significate in our success. A special mention goes to our friend, whose unwavering support and encouragement have been a constant source of motivation throughout the project. Finally, we would like to express our thanks to all faculty members of computer science and engineering department for their valuable suggestion an input during the project's duration. Their expertise has greatly contributed to our project's development.

AYUSH KUMAR   **KM MINAKSHI**

(1903700100003)   **(2003700109003)**

SHUBHAM   **AAYUSHI**

(2003700109005)   **(2003700109001)**

TABEL OF CONTENT

(<u>INDEX</u>)

LIST OF FIGURES

LIST OF TABELS

INTRODUCTION :

Title of the project :- Women safety sentiment analysis based on twitter data.

There are certain types of harassment and Violence that are very aggressive including staring and passing comments and these unacceptable practices are usually seen as a normal part of the urban life. There have been several studies that have been conducted in cities across India and women report similar type of sexual harassment and passing off comments by other unknown people. The study that was conducted across most popular Metropolitan cities of India including Delhi, Mumbai and Pune, it was shown that 60 % of the women feel unsafe while going out to work or while travelling in public transport. Women have the right to the city which means that they can go freely whenever they want whether it be too an Educational Institute, or any other place women want to go. But women feel that they are unsafe in places like malls, shopping malls on their way to their job location because of the several unknown Eyes body shaming and harassing these women point Safety or lack of concrete consequences in the life of women is the main reason of harassment of girls. There are instances when the harassment of girls was done by their neighbors while they were on the way to school or there was a lack of safety that created a sense of fear in the minds of small girls who throughout their lifetime suffer due to that one instance that happened in their lives where they were forced to do something unacceptable or was sexually harassed by one of their own neighbor or any other unknown person. Safest cities approach women safety from a perspective of women rights to the affect the city without fear of violence or sexual harassment. Rather than imposing restrictions on women that society usually imposes it is the duty of society to imprecise the need of protection of women and also recognizes that women and girls also have a right same as men have to be safe in the City.

Analysis of twitter texts collection also includes the name of people and name of women who stand up against sexual harassment and unethical behavior of men in Indian cities which make them uncomfortable to walk freely. The data set that was obtained through Twitter about the status of women safety in Indian society was for processed through machine learning algorithms for the purpose of smoothening the data by removing zero values and using Laplace and porter's theory is to developer method of analyzation of data and remove retweet and redundant data

from the data set that is obtained so that a clear and original view of safety status of women in Indian society is obtained .
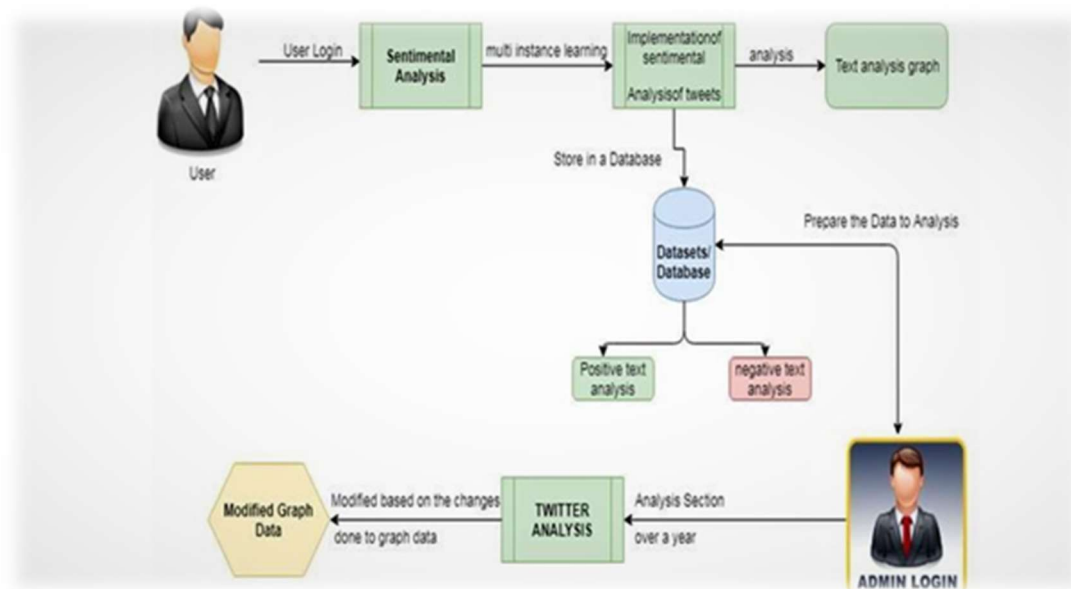

Fig 1.1

Twitter in this modern era has emerged as a ultimate microblogging social network consisting over hundred million and generates over five hundred messages known as 'tweets' everyday. Twitter with such a massive audience has magnetized users to emit their perspective and judgmental about every existing issue and topic of internet, therefore twitter is an informative source for all the zones like institutions, companies, and organizations. On the twitter, users will share their opinions and perspective in the tweets section. This tweet can only contain 140 characters, thus making the users to compact their messages with the help of abbreviations, slang, shot forms, emoticons, etc. In addition to this, many people express their opinions by using polysemy and sarcasm also.

Keywords:

Generic Technology Keywords :

Women , safety , sexual harassment, hashtag, sentiment analysis, programming, coding.

Technology keywords:

Python programming language, Ide : VS Code, Machine learning.

1.2 **OBJECTIVE AND SCOPE**

Objective:

The objective of the project "Women Safety Analysis According to Twitter Data Using Machine Learning" is to utilize machine learning techniques to analyze Twitter data and gain insights into the safety concerns and issues faced by women. The project aims to identify patterns, trends, and potential risks associated with women's safety on the platform, with the ultimate goal of developing strategies and interventions to address these concerns.

Scope:

The scope of the project encompasses the following key areas:

1. Data Collection: The project will involve collecting a substantial amount of Twitter data related to women's safety. This may include tweets, user profiles, hashtags, and other relevant information. The collection process should adhere to ethical guidelines and respect user privacy.

2. Preprocessing and Cleaning: The collected Twitter data will undergo preprocessing and cleaning to remove noise, irrelevant content, duplicates, and spam. Natural Language Processing (NLP) techniques will be applied to extract meaningful features and prepare the data for analysis.

3. Sentiment Analysis: Machine learning algorithms will be employed to perform sentiment analysis on the Twitter data. This will help determine the overall sentiment (positive, negative, or neutral) associated with women's safety-related discussions.

4. Topic Modeling: The project will utilize topic modeling techniques such as Latent Dirichlet Allocation (LDA) or similar algorithms to identify prevalent topics and themes related to women's safety. This will enable the identification of key concerns and issues faced by women on Twitter.

5. Risk Identification: Machine learning models will be developed to identify potential risks and threats to women's safety based on the analyzed Twitter data. This may involve identifying patterns of harassment, abuse, or discriminatory behavior.

6. Visualization and Reporting: The project will involve visualizing the analyzed data and generating meaningful reports and insights. Data visualization techniques such as graphs, charts, and word clouds will be utilized to present the findings effectively.

7. Recommendations and Interventions: Based on the analysis and insights gained from the project, recommendations and interventions will be proposed to address the identified safety concerns. These recommendations may include policy suggestions, user education initiatives, or technological interventions.

It's important to note that the project's scope is limited to analyzing Twitter data, and the findings may not represent the entire spectrum of women's safety concerns. The project does not involve real-time monitoring or direct intervention on the platform but aims to contribute valuable insights for addressing women's safety issues in the online space.

REQUIREMENT ANALYSIS :

Requirements Analysis for the project "Women Safety Analysis According to Twitter Data Using Machine Learning":

2.1. Functional Requirements:

2.1.1 Data Collection: - The system should have an CSV file , which is collection of tweets from the twitter or utilize a reliable data scraping tool to collect Twitter data related to women's safety.

- The system should ensure the collection of diverse data, including tweets, user profiles, hashtags, and contextual information.

2.1.2 Data Preprocessing: - The system should implement data preprocessing techniques to clean and normalize the collected Twitter data.

- Noise, irrelevant content, duplicates, and spam should be removed from the data.

- Techniques such as tokenization, stemming, and lemmatization should be applied to prepare the data for analysis.

2.1.3 Natural Language Processing (NLP): - The system should employ NLP techniques to extract meaningful features from the Twitter data.

- Sentiment analysis, entity recognition, part-of-speech tagging, and named entity recognition should be performed to analyze the data.

2.1.4 Machine Learning Algorithms: - The system should support various machine learning algorithms suitable for analyzing Twitter data.

- Supervised learning algorithms, such as classification or regression, should be available for risk identification.

- Unsupervised learning algorithms, like clustering or topic modeling, should be provided for theme identification.

- Deep learning models should be supported for advanced analysis tasks.

2.1.5 Model Training and Evaluation :- The system should facilitate the training of machine learning models using labeled or annotated datasets related to women's safety.

 - Evaluation metrics should be incorporated to assess the performance and accuracy of the trained models.

2.1.6 Visualization: - The system should provide data visualization capabilities to present the analyzed data in a visually appealing and understandable manner.

 - Charts, graphs, heatmaps, and word clouds should be generated to represent patterns, trends, and key findings.

2.2. Non-Functional Requirements:

2.2.1 Scalability and Performance:- The system should be capable of handling large-scale Twitter data efficiently.

 - It should utilize computing resources effectively to ensure reasonable processing times.

2.2.2 Privacy and Ethical Considerations: - The system should adhere to ethical guidelines while collecting and analyzing Twitter data, respecting user privacy and consent.

 - Measures should be implemented to handle sensitive information and avoid biases or perpetuation of harm.

2.2.3 User-Friendly Interface: - The system should have an intuitive and user-friendly interface for ease of use.

- Users should be able to interact with the system to define analysis parameters and explore results effectively.

2.2.4 Accuracy and Reliability: - The system should aim to provide accurate and reliable analysis results.

- The implemented machine learning algorithms should be robust and validated against appropriate evaluation metrics.

2.2.5 Integration and Compatibility: - The system should be compatible with relevant tools and technologies commonly used in the field of machine learning and data analysis.

 - Integration with external data sources, such as crime statistics or demographic information, should be facilitated.

2.2.6 Hardware and Software : - For this project implementation we need PCs or Laptops for implementing the code of project.

-Hardware system have Minimum 4 GB of RAM , 512 GB HDD or 256 GB SSD , and minimum i3 or Ryzen 3 3200U processor.

-For software it need Window Operating System Window 10 or Window 11.

2.2.7 Documentation and Support: - The system should have comprehensive documentation, including installation instructions, usage guidelines, and explanations of implemented algorithms.

 - Adequate support and troubleshooting resources should be provided to assist users in utilizing the system effectively.

Meeting these requirements will ensure the successful implementation of the project, enabling in-depth analysis of women's safety concerns on Twitter using machine learning techniques.

FEASIBILITY STUDY

Feasibility Study of the project "Women Safety Analysis According to Twitter Data Using Machine Learning":

3.1. Technical Feasibility:

  - Availability of Necessary Tools and Technologies: Assess the availability of suitable tools, libraries, and frameworks for data collection, preprocessing, machine learning, and visualization of Twitter data.

  - Computational Resources: Determine if the required computing resources, such as servers or cloud-based platforms, are accessible to handle the scale of data and computational demands involved in the project.

  - Expertise and Skillset: Evaluate the technical expertise and skills required to implement the project, including proficiency in machine learning, natural language processing, data preprocessing, and data visualization.

3.2. Data Feasibility:

  - Availability of Relevant Twitter Data: Investigate the availability and accessibility of Twitter data related to women's safety. Ensure that a substantial amount of diverse and representative data can be obtained for analysis.

  - Quality and Completeness of Data: Assess the quality and completeness of the collected data, considering factors such as data noise, spam, duplicates, and any limitations or biases inherent in the Twitter data.

3.3. Ethical and Legal Feasibility:

  - Data Privacy and Consent: Ensure compliance with ethical guidelines, privacy regulations, and terms of service provided by Twitter. Verify that user privacy and consent are respected throughout the data collection and analysis process.

- Compliance with Data Protection Laws: Assess the project's compliance with relevant data protection laws and regulations, particularly regarding the handling and storage of personal data.

3.4. Time Feasibility:

- Project Timeline: Evaluate the estimated time required for each phase of the project, including data collection, preprocessing, analysis, model training, and visualization. Consider any potential delays or challenges that may impact the project's timeline.

3.5. Financial Feasibility:

- Cost Analysis: Conduct a cost analysis of the required resources, including computing resources, software licenses, and any additional tools or services needed for the project's implementation.

- Budget Availability: Determine the availability of budgetary resources to cover the project costs, including any potential expenses for data acquisition, computing resources, and expertise.

3.6. Operational Feasibility:

- User Acceptance: Assess the stakeholders' willingness and readiness to adopt and utilize the project's outcomes and recommendations, such as policymakers, social activists, or organizations working on women's safety issues.

- Integration with Existing Systems: Consider the feasibility of integrating the project's findings and recommendations with existing systems, policies, or initiatives addressing women's safety.

RELATED WORK

Sentiment Analysis (SA) is a subject of study that investigates people's sentiments, views, assessments, appraisals, attitudes and emotions in the direction of entities such as individuals, services, organizations, issues, products, topics and their characteristics. It is also known as opinion mining, sentiment mining, subjectivity analysis, review mining, opinion extraction, emotion analysis, etc. Furthermost of the prevailing approaches have used the terms sentiment analysis and opinion mining interchangeably. According to the research Mathematics, this mining is defined as a quintuple. Sentiment mining = (t, s, h, T) Where 't' is the target opinion, 's' is the sentiment about 't', 'h' is the holder opinion and 'T' is the time. There are 3Approaches in Sentiment Analysis. They are: 1. Machine-learning approach, 2. Lexicon based approach, 3. Hybrid approach.

4.1. Table: Represents the sentimental classification methods

| Sentiment methods	Classification	Pros and Cons
Lexicon based	Dictionary based. Corpus based. Ensemble approaches	Pros: Best for domain reliant on, larger-term coverage. Cons: Only Finite number of words in the lexicons
Machine learning Based	Support vector machines. Bayesian networks. Naïve Bayes. Random forest.	Pros: Capacity to adjust and make prepared models for explicit purposes and settings. Cons: Low relevance for new information, since it is important of marked information.
Hybrid based	Lexicon and machine learning based.	Pros: High exactness of new information. Slant vocabulary developed utilizing public assets for assumption discovery. Notion words as highlights in the AI technique. Cons: Noisy data.

Sentiment analysis approach:
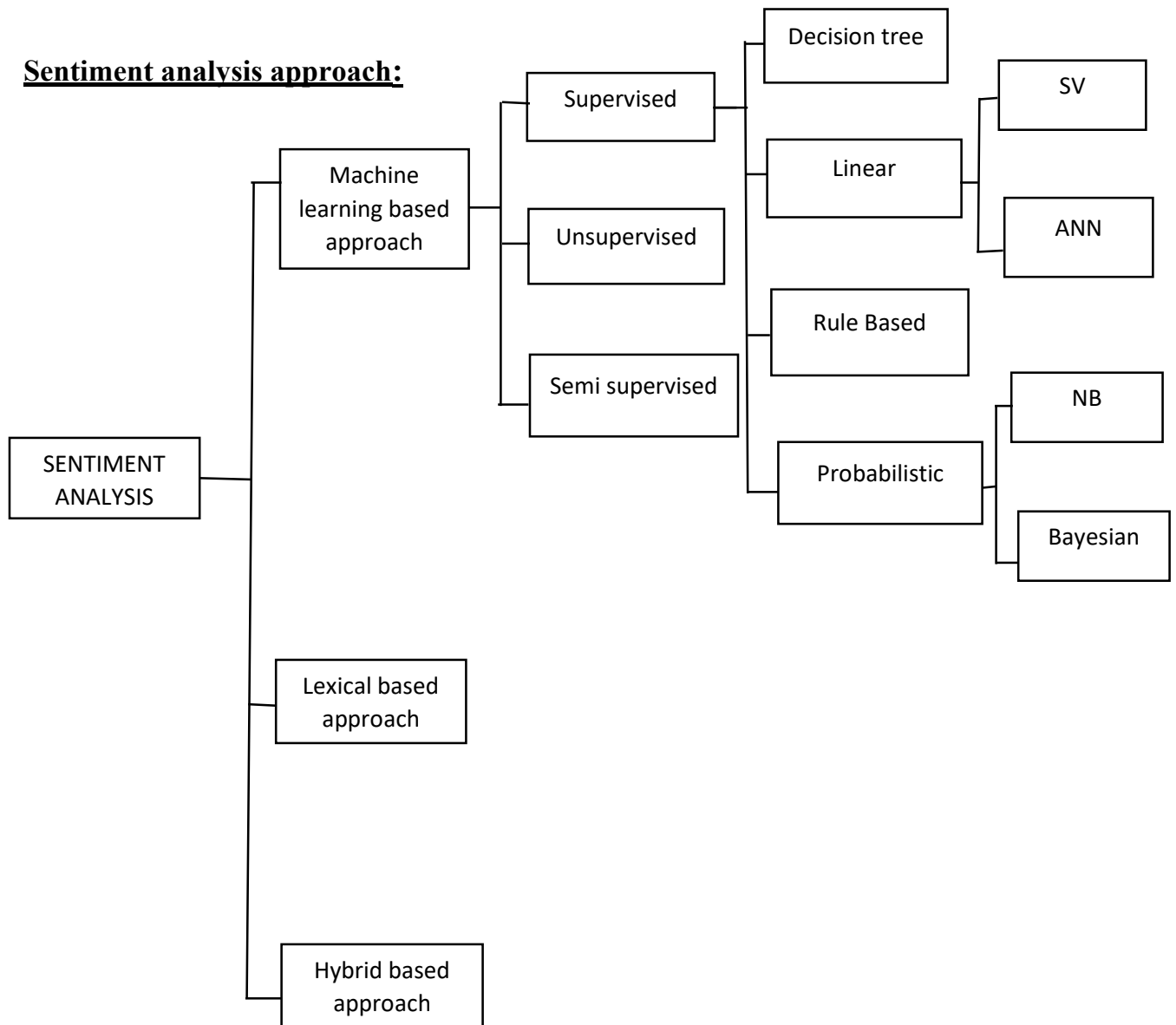
Fig.4.2.Sentiment analysis approach

4.3. Table: Different Literature on Sentiment Analysis

Authors	Corpus	Methods	Comments
Vanzo et al (2014)	Twitter	SVM	Markovian design for context based polarization detection
Shoukry and Rafea (2012)	Twitter	SVM,NB	Sentence level sentiment analyze for Arabic
Duwairi and Qarqaz (2014)	Twitter ,FB	Multiple	Rapid miner for text classification on social media reviews
Amolik et al (2016)	Twitter	SVM,NB	movie reviews using machine learning
Kumar et.al 2019	Twitter	SVM	Content based, Bag of words polarity check on 3 cities data

OUR METHODOLOGY:
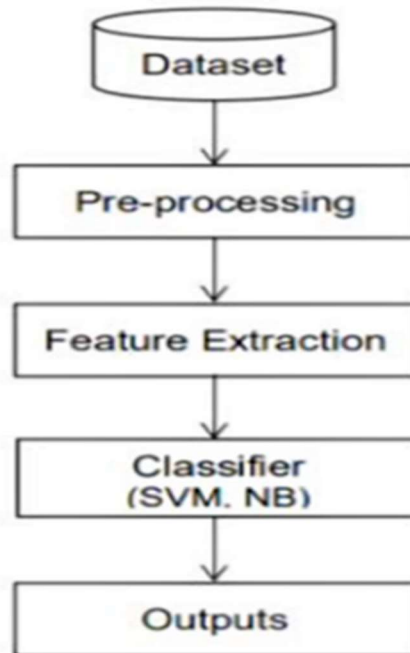
Flow diagram of our methodology:

Fig 5.1 Flow diagram of our methodology

In the Sentiment Analysis the following steps are major to identify the positive ,negative

or neutral of the twitter post. They are:

5.2 .Collection of dataset :

Utilize Twitter's API or a reliable data scraping tool to collect Twitter data related to women's safety.

Define relevant search queries, hashtags, and filters to ensure the collection of diverse and representative data.

Retrieve tweets, user profiles, hashtags, and contextual information for analysis.

5.3 Pre- Processing the Dataset :

Clean the collected data by removing noise, duplicates, and irrelevant content.

Perform data cleaning techniques such as removing special characters, URLs, and emojis.

Apply tokenization, stemming, and lemmatization to normalize the text data.

Handle missing data, perform data imputation if necessary, and ensure data consistency.

5.4 Feature Extraction:

Algorithm- Bag of words:

In this work, we used Bag of Words to extract features from text documents.

After extraction, these features used for training machine learning algorithms. It

makes a jargon of the apparent multitude of novel words happening in all the

reports in the preparation set. Bag of words features containing term frequencies

of each word in each document, i.e. the number of occurrence and not sequence

or order of words matters. This can be done by Count Vectorizer method in

Python.

5.5 Classification:

Machine learning approach:

classification problem is applied if the output variable is a label or category,

such as "Rainy" or "Sunny" or "disease" and "no disease" or in our work

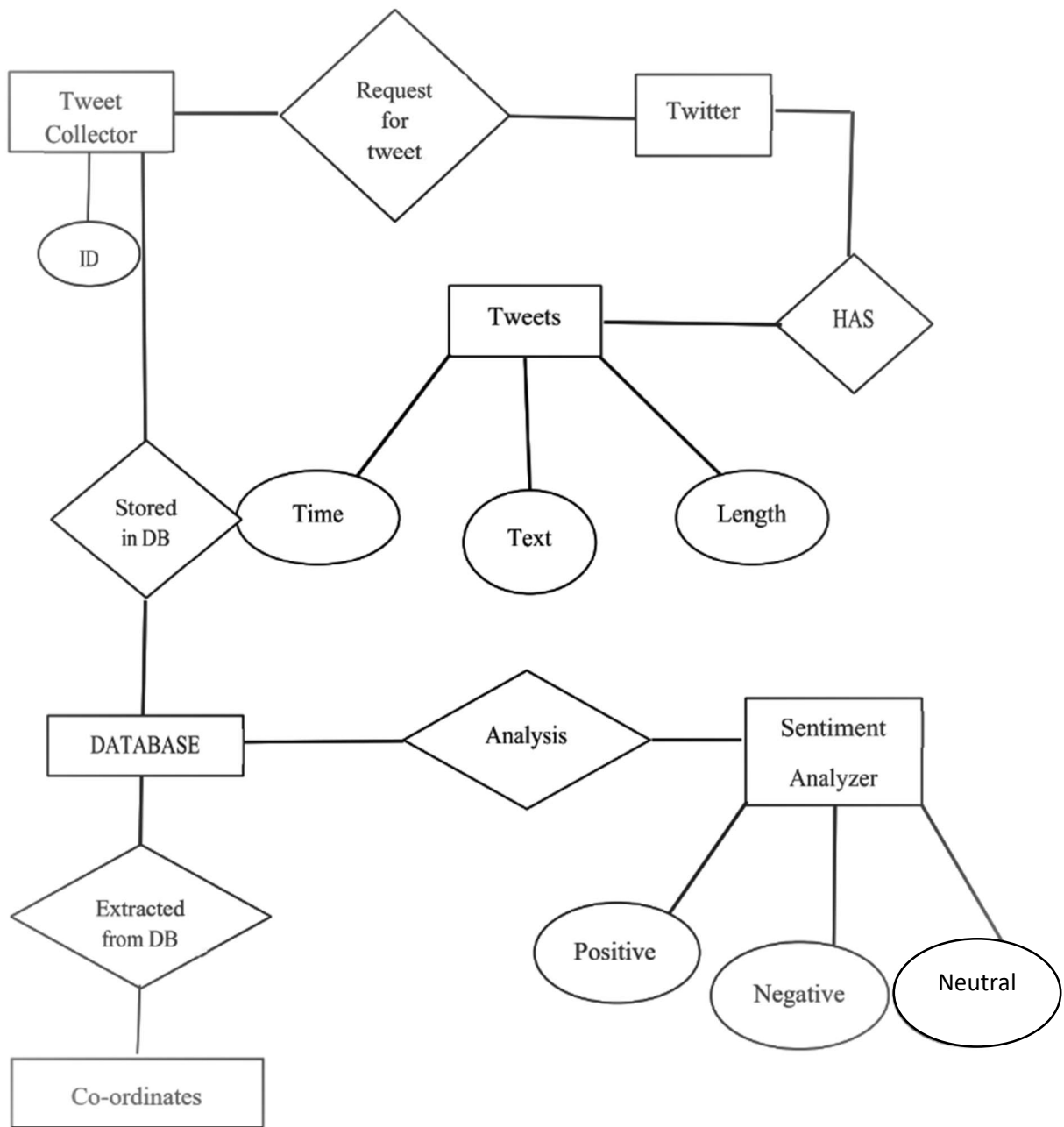"Positive" or "Negative".

5.6. ER Diagram:

Fig.5.2. ER Diagram:

CODING :

Language: We use Python programming language in our project because Python has lots of libraries which are used for Data Analytics , Data science, Machine learning etc.

Python libraries are open source and free of cost.

CODE:

```
import tkinter as tk

from tkinter import *

from tkinter.font import BOLD, Font

from tkinter.filedialog import askopenfilename, asksaveasfilename

import pandas as pd

from tkinter import filedialog

from tkinter import messagebox

from sklearn.feature_extraction.text import CountVectorizer

from sklearn.model_selection import train_test_split

from sklearn.naive_bayes import MultinomialNB

from sklearn.svm import SVC

import matplotlib.pyplot as plt



## Create the DynamicGUI class for User Interface use Tkinter library.



class DynamicGUI(tk.Frame):
```

```python
def __init__(self,master=None):

    super().__init__(master)

    self.create_widgets()

    self.create_buttons()

    self.text()

    self.master = master
```

Creating the function of WIdgets for GUI .

```python
def create_widgets(self):

    self.bold25 = Font(self.master, size=25, weight=BOLD,underline=True,slant="italic")

    self.bold20 = Font(self.master,family= "Times", size=12,
weight=BOLD,underline=True,slant="italic")

    self.label1 = tk.Label(self.master,height=3,width=90, bg='aquamarine1',text="WOMEN
SAFETY ANALYSIS ACCODING TO TWITTER DATA",compound='center',font=self.bold25)

    self.label1.pack(ipadx=100)
```

Create text box for user interface.

```python
def text(self):

    self.textbox = tk.Text(self.master)

    self.textbox.pack(padx=10,pady=0,ipadx=800,ipady=1000)

    Scroll= tk.Scrollbar(self.textbox)

    Scroll.pack(side =RIGHT, fill=Y)

    Scroll.config(command=self.textbox.yview)

    self.textbox.config(yscrollcommand=Scroll.set)
```

##Creating buttons for GUI with tkinter .

```python
def create_buttons(self):

    self.label=tk.Label(self.master,text="options",bg='red',bd='17 pixel'
,width=200,height=8)

    self.label.place(x=5,y=40)


    self.button_frame=tk.Frame(self.label,height=5,width=2)

    self.button_frame.place(x=50,y=70)
```

```python
        self.button_frame.pack()


        self.button = tk.Button(self.button_frame, text="Upload and Read tweets
data",font=self.bold20,bg="dark slate grey", fg="white",
command=lambda:upload_data(self.textbox))

        self.button.place(x=50, y=50)


        self.button.pack(side="left",padx=15,pady=5,ipadx=15,ipady=10)

        self.button_frame.pack(ipadx=10,ipady=5)

        self.label.pack()


        # self.button = tk.Button(self.button_frame, text="Read tweets",font=self.bold20,
bg="dark slate grey", fg="white", command=read_tweets)

        # self.button.pack(side="left",padx=15,pady=5,ipadx=15,ipady=10)


        self.button = tk.Button(self.button_frame, text="Tweet cleaning",font=self.bold20,
bg="dark slate grey", fg="white", command=lambda: tweet_cleaning(self.textbox))

        self.button.pack(side="left",padx=15,pady=5,ipadx=15,ipady=10)


        self.button = tk.Button(self.button_frame, text="Run machine learning algorithm",
font=self.bold20, bg="dark slate grey", fg="white", command=lambda:
run_algorithm(self.textbox))

        self.button.pack(side="left",padx=15,pady=5,ipadx=15,ipady=10)
```

```python
        self.button = tk.Button(self.button_frame, text="Analysis graph",font=self.bold20,
bg="dark slate grey", fg="white", command=lambda: analysis_graph(self.textbox))

        self.button.pack(side="left",padx=15,pady=5,ipadx=15,ipady=10)



        self.label.pack(ipadx=20,ipady=5,padx=10,pady=15)
```

##Creating the function for the module upload data .

```python
def upload_data(text_widget):

    file_path = filedialog.askopenfilename(filetypes=[("CSV files", "*.csv")])

    if file_path:

      try:

        df = pd.read_csv(file_path)

        messagebox.showinfo("Success", "Dataset uploaded successfully!")

        text_widget.insert(tk.END, df.to_string())  # Display the dataframe in the textbox

      except:

        messagebox.showerror("Error", "Failed to upload the dataset!")
```

Creating the function of Tweet Cleaning module.

```python
def tweet_cleaning(text_widget):

    text_data = text_widget.get("1.0", "end-1c")  # Get the content of the textbox


    # Clean duplicates

    unique_tweets = list(set(text_data.split("\n")))

    cleaned_data = "\n".join(unique_tweets)


    # Clean null data

    cleaned_data = cleaned_data.replace("null", "")


    # Clean incomplete data


    cleaned_data = cleaned_data.replace("...", "")


    # Clean special characters and unnecessary data


    special_chars = ['!', '@', '#', '$', '%', '^', '&', '*', '(', ')', '_', '+', '=', '[', ']', '{', '}', '|',

            '\\', ':', ';', '<', '>', ',', '.', '?', '/', '~', '`']

    for char in special_chars:

        cleaned_data = cleaned_data.replace(char, "")

    ##arrange data in ascending format
```

```python
        cleaned_data = "\n".join(sorted(cleaned_data.split("\n"), key=str.lower))

        text_widget.delete("1.0", "end")  # Clear the existing content

        text_widget.insert("1.0", cleaned_data)  # Insert the cleaned and sorted data
```

Creating the function of Run Algorithm Module.

```python
def run_algorithm(text_widget):


    pass
```

Creating the function of Analysis Graph Module.

```python
def analysis_graph(text_widget):

    pass

root = tk.Tk()

root.geometry('1030x500')

# if __name__ == "__main__":

app = DynamicGUI(master=root)

app.mainloop()
```

1.Home screen.

This is Graphical user Interface (GUI) of our project with Button (Upload and read tweets, tweet cleaning, run machine learning algorithm and Analysis graph)

I
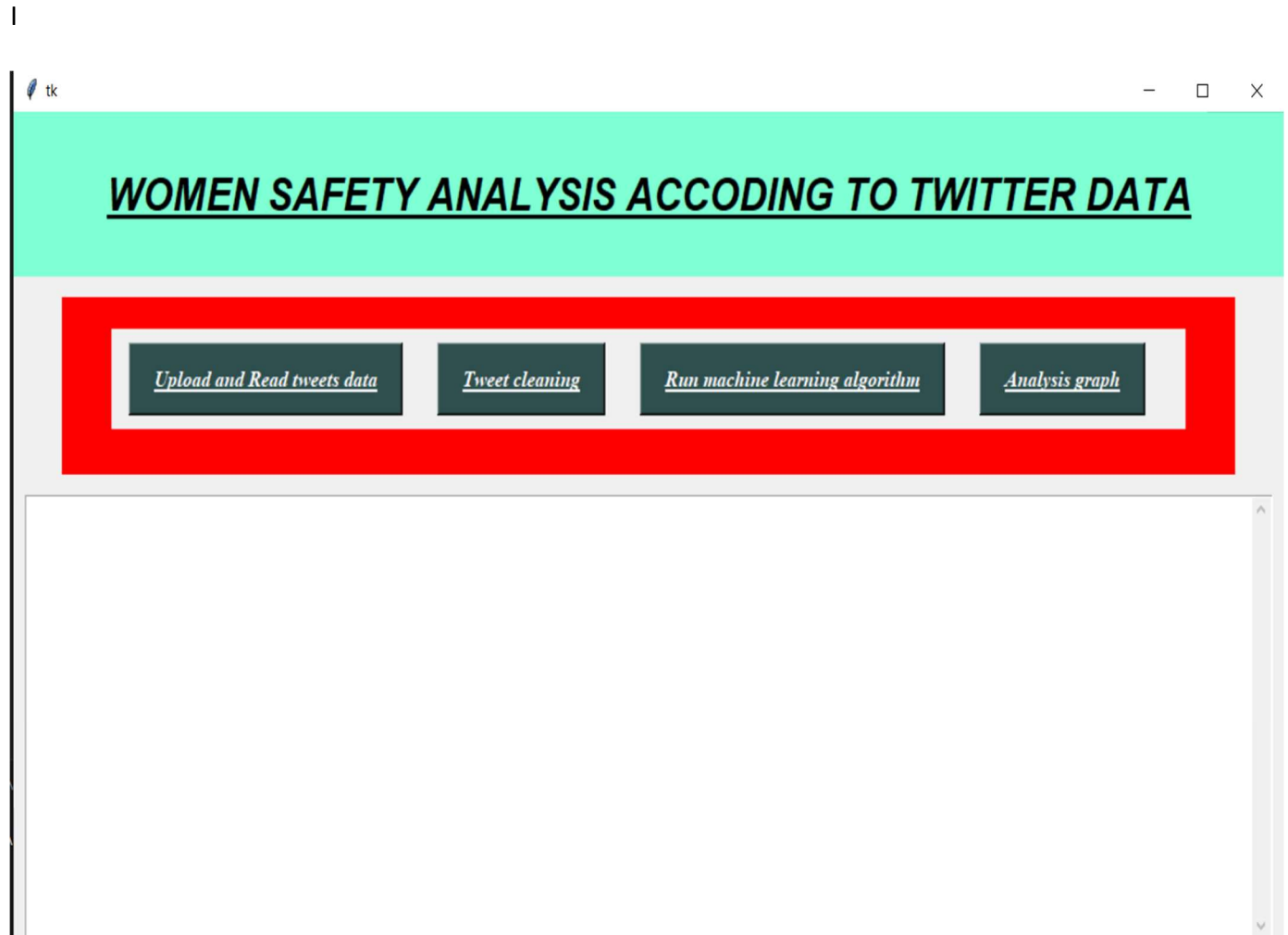
Fig. 7.1 : Home Screen

1. Selecting the Data Set for Process:

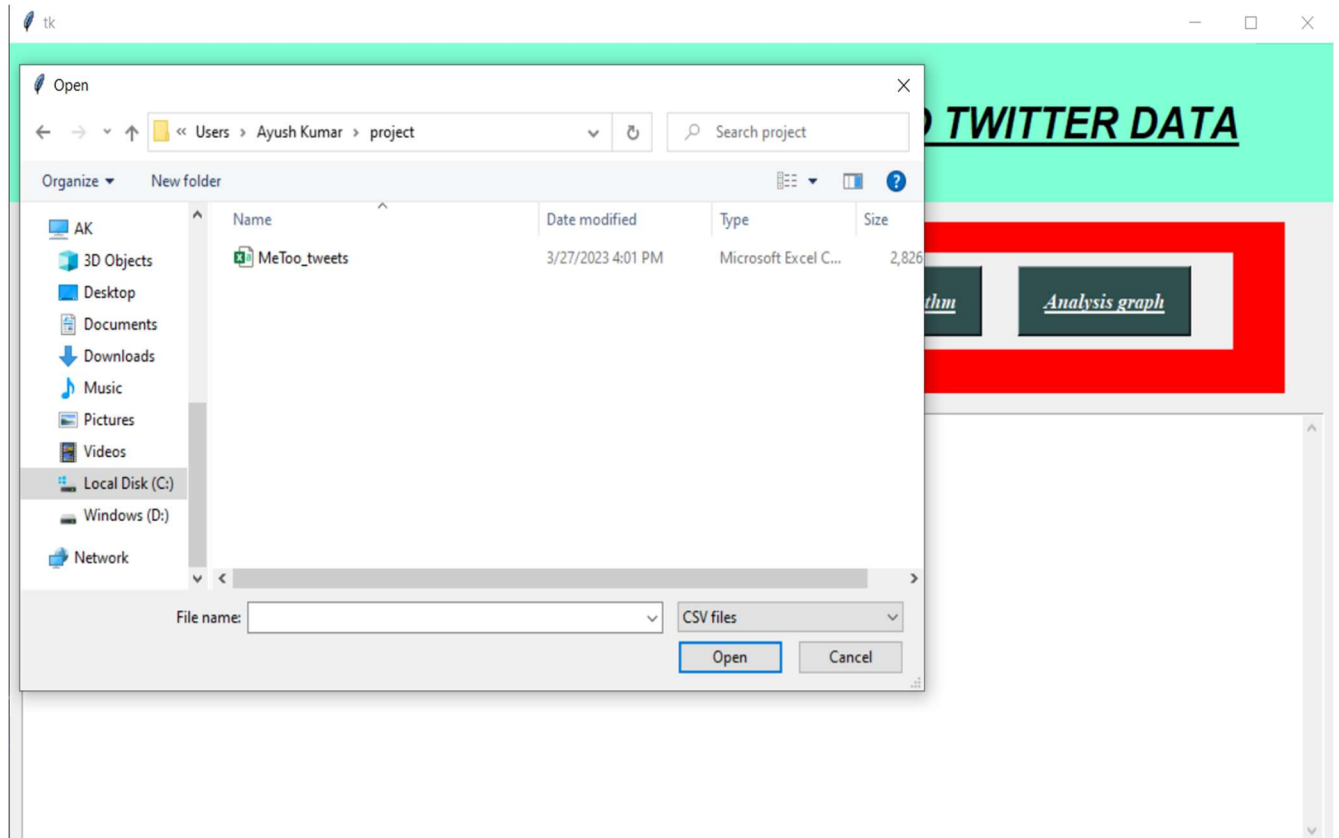With the help of upload and tweet button we select the data set and upload it in application.

Fig.7.2 : Select file

2. Upload successfully:

When we select the data file and upload it .Then it shows a message "Dataset Uploaded Successfully. If the data is not uploaded the it shows a message "Dataset Uploaded Failed".
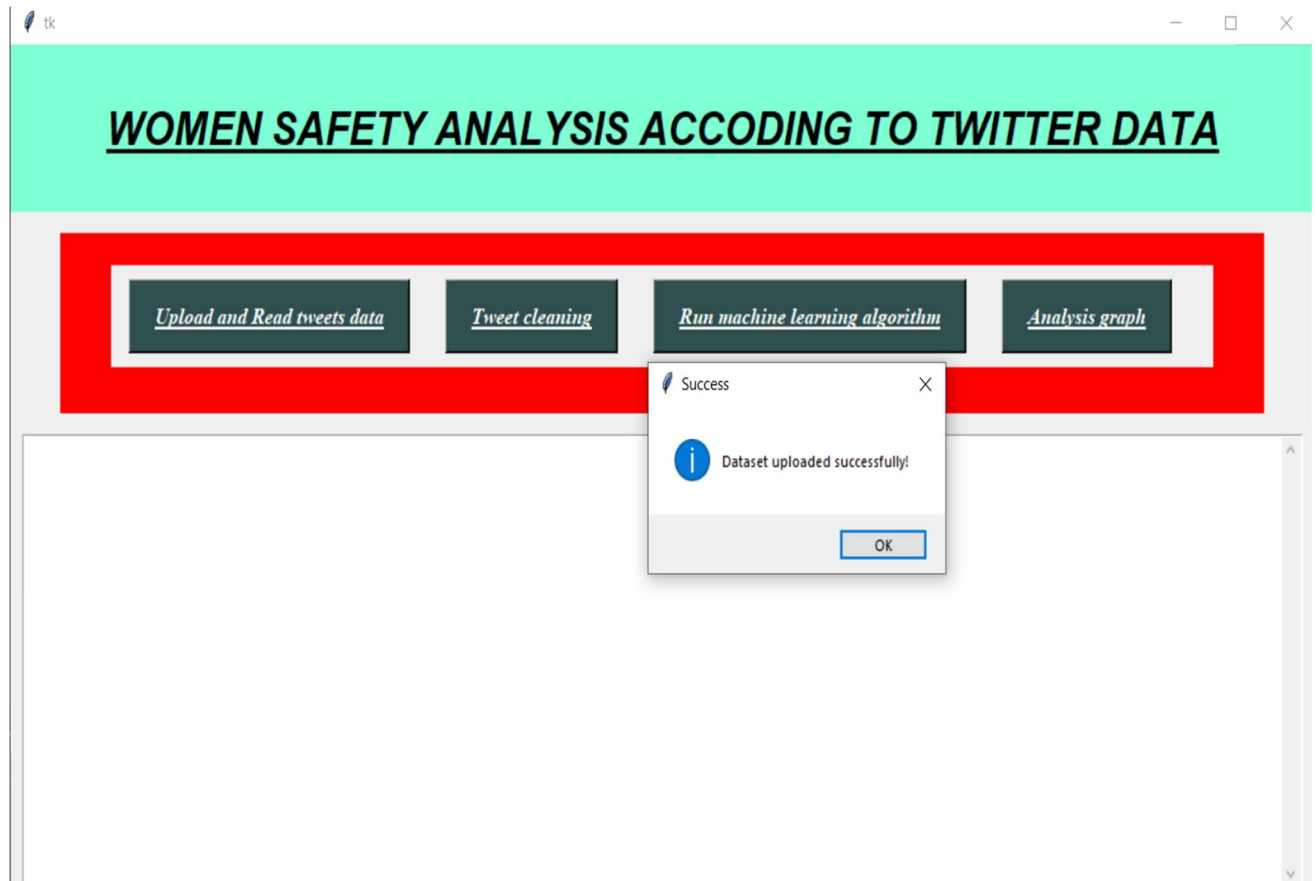

Fig. 7.3: Uploaded successfully

3. Upload and read data:

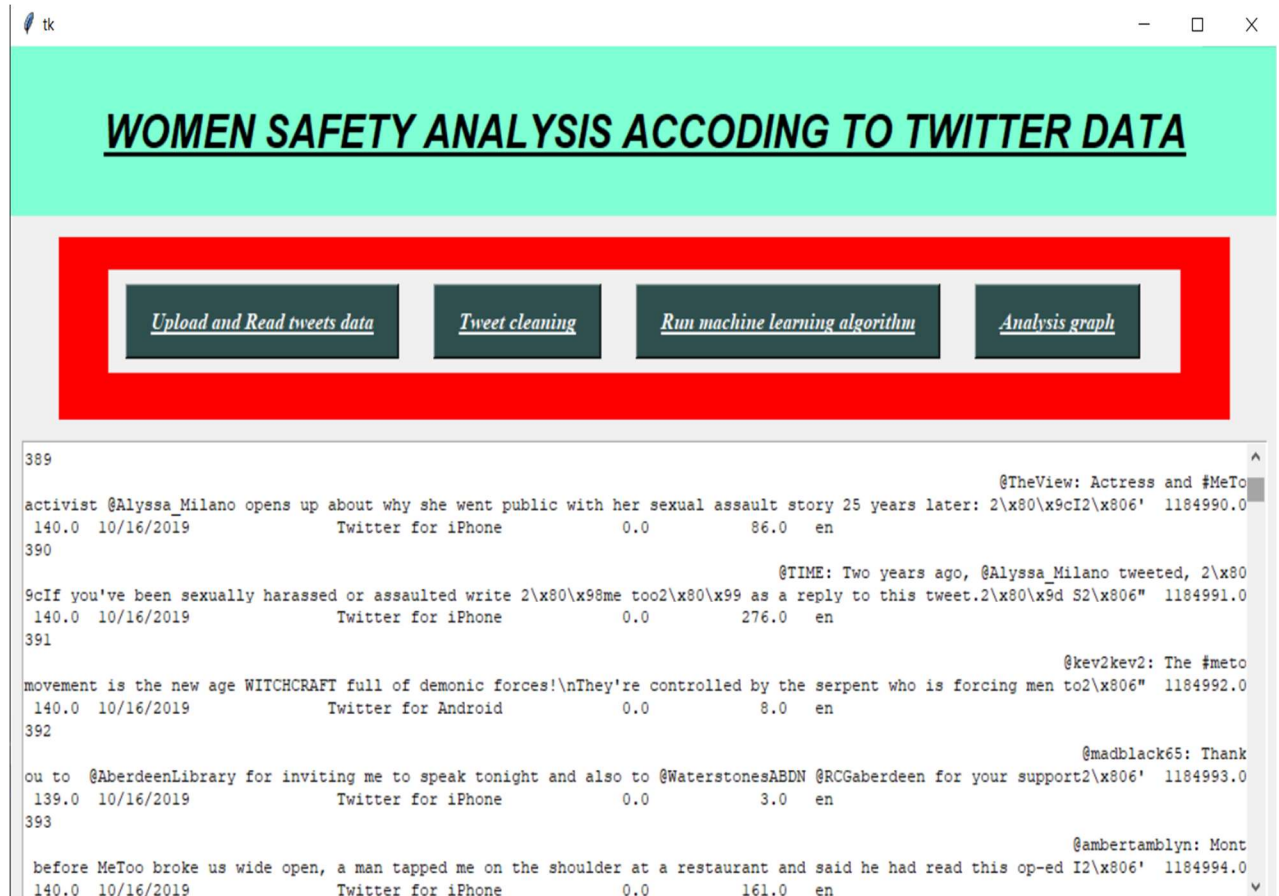After the loading data, We read the data set which is uploaded

Fig.7.4: Upload and Read

4. Tweet cleaning:

After upload and read the data , We clean the tweet .Our data set has lots of tweets some are incomplete tweets, duplicate, empty, undefined or some with the special character . So we'll go for tweet cleaning.
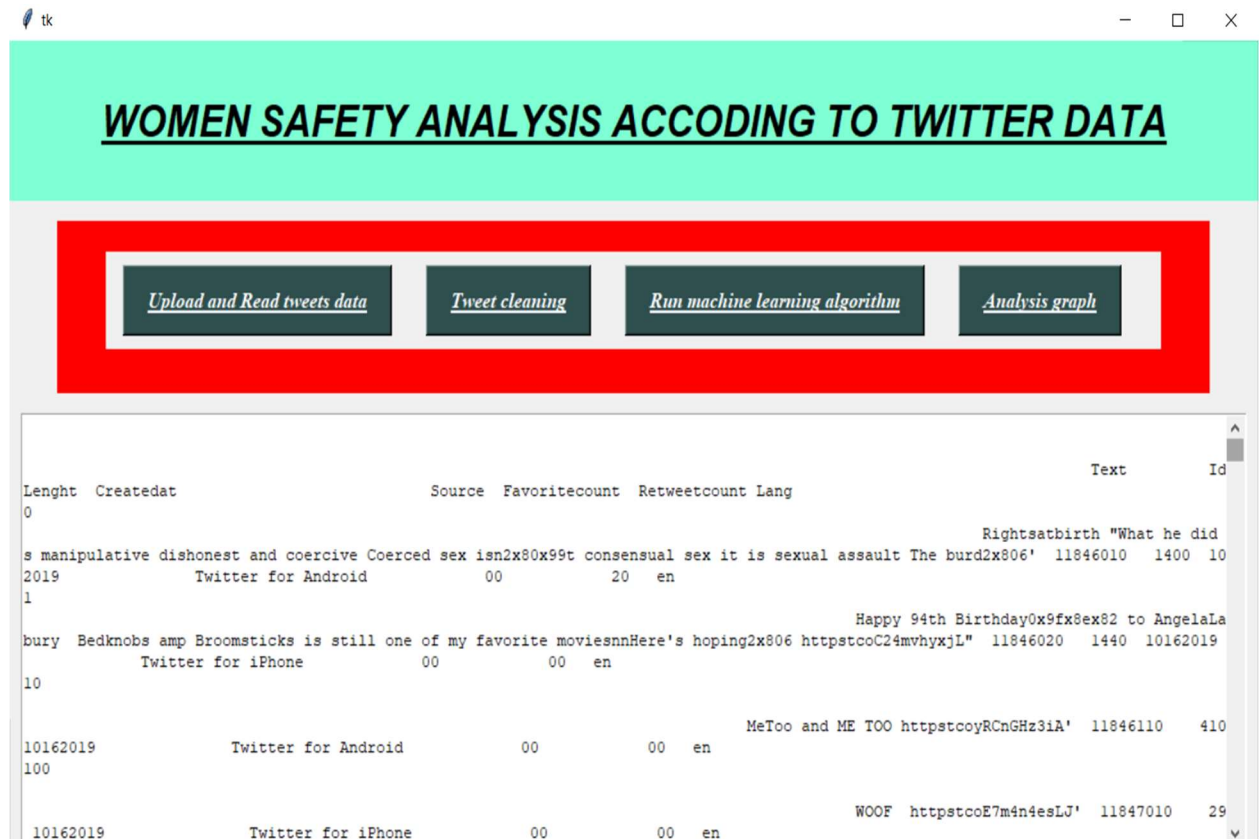

Fig.7.5: Tweet cleaning

5. Run Machine learning algorithm:

After the tweet cleaning, We apply the Machine learning algorithm for sentiment analysis on the tweets.
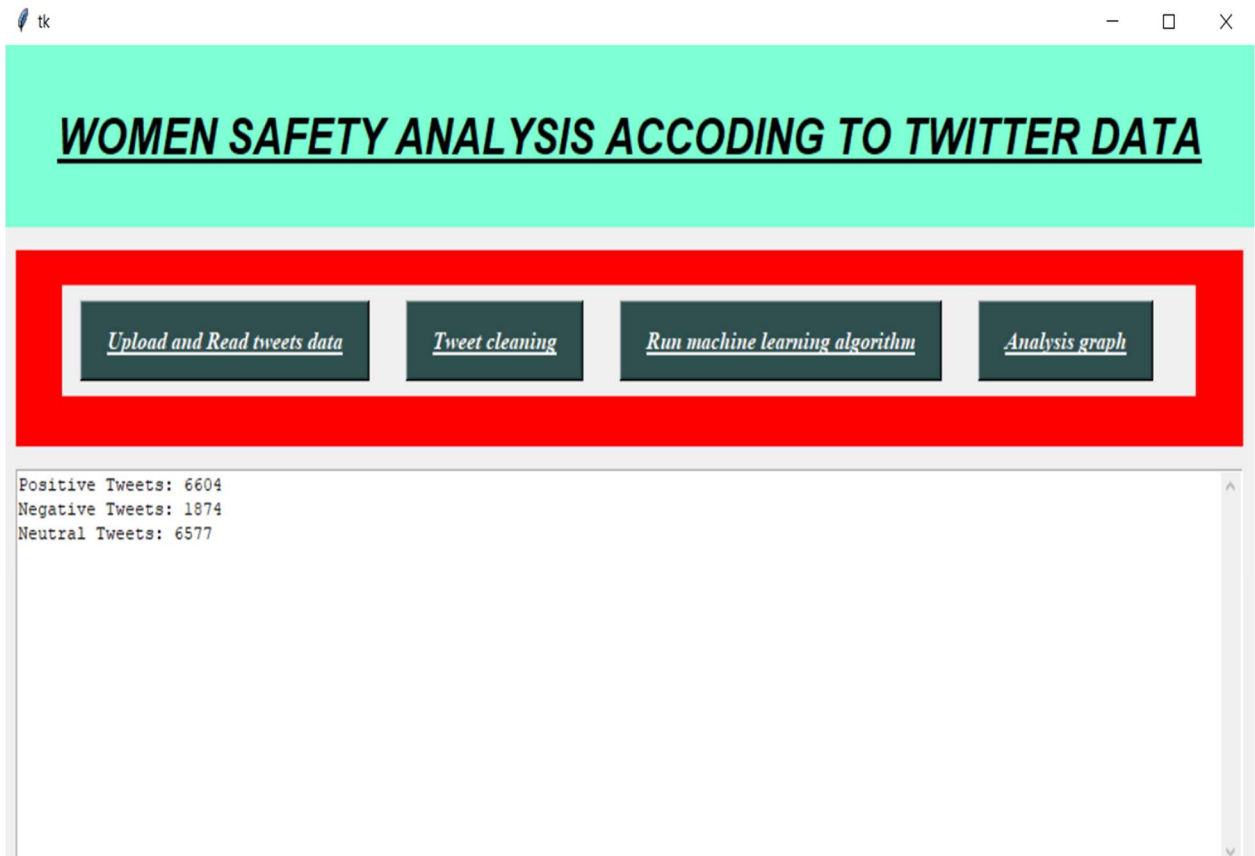

Fig 7.6: Run Algorithm

8. Analysis graph:

After the machine learning algorithm we get our analysis result in the form of graph (basically in Pie chart form)
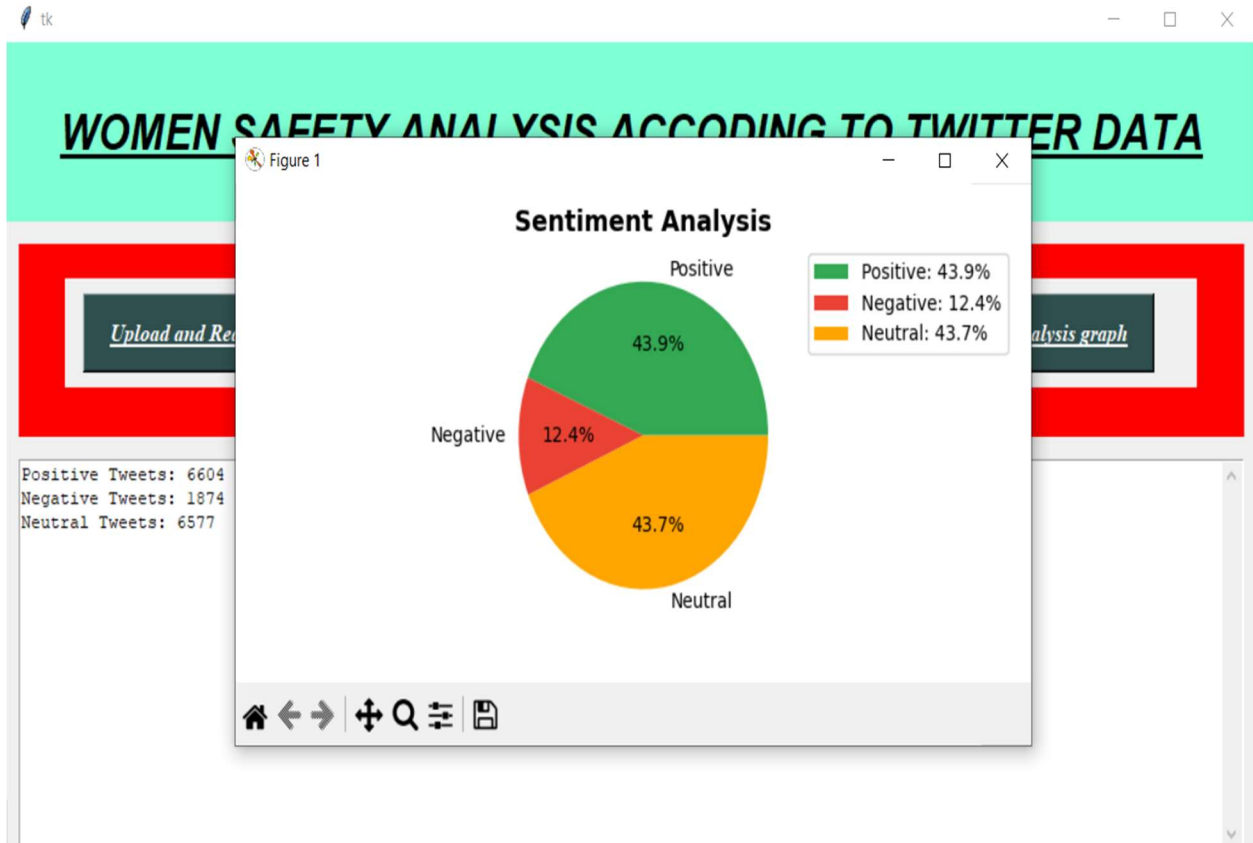

Fig.7.7: Result Graph

CONCLUSION:

In conclusion, the project "Women Safety Analysis According to Twitter Data Using Machine Learning" showcases the power of leveraging data-driven approaches to drive positive change in women's safety. By harnessing the capabilities of machine learning and analysis of Twitter data, this project contributes to creating a safer and more inclusive digital space for women. Continued research, collaboration, and innovation in this field will lead to greater strides in ensuring women's safety and empowerment in the online world.

FUTURE SCOPE:

For the future enhancement, we can extend to apply these machine learning algorithms on different social media platforms like facebook and instagram also since in our project only twitter is considered. Present ideology which is proposed can be integrated with the twitter application interface to reach larger extent and apply sentimental analysis on millions of tweet to provide more safety.

REFERENCES:

[1] Eugene Charniak and additionally Mark Johnson. "Coarse-to-exceptional nbest parsing and MaxEnt discriminative reranking." Process of the 43rd annual assembly on association for computational linguistics. Organization for Computational Grammar, 2005.

[2] Gupta B, Negi M, Vishwakarma K, Rawat G &Badhani P (2017). "Research of Twitter sentiment evaluation making use of system mastering algorithms on Python." International Journal of Computer System Applications, one hundred sixty-five(nine) 0975-8887.

[3] Adam Bermingham and Alan F. Smeaton. "Classifying sentiment in microblogs: is brevity an advantage?." complaints of the 19th ACM global conference on information and expertise control. ACM, 2010.

[4] Mamgain N, Mehta E, Mittal A & Bhatt G (2016, March). "Belief evaluation of top schools in India making use of Twitter data." In Computational Strategies, in Details and Interaction Technologies (ICCTICT).

[5] Soo-Min Kim and Eduard Hovy. "determining the sentiment of evaluations." court cases of the twentieth international conference on Computational Linguistics. association for Computational Linguistics, 2004.

[6] Dan Klein as well as Christopher D. Manning. "Accurate unlexicalized parsing." Procedures of the forty first Yearly Meeting on Organization for Computational LinguisticsVolume 1. Association for Computational Linguistics, 2003

[7] Revathy K & Sathiyabhama B. (2013). A hybrid approach for supervised twitter sentiment classification. International Journal of Computer Science and Business Informatics.

[8] Sahayak V, Shete V & Pathan A (2015). "Sentiment analysis on twitter statistics." international journal of revolutionary research in advanced Engineering (IJIRAE), 2(1), 178-183.