

Sentiment Analysis Based on News Topics during COVID-19 using Twitter

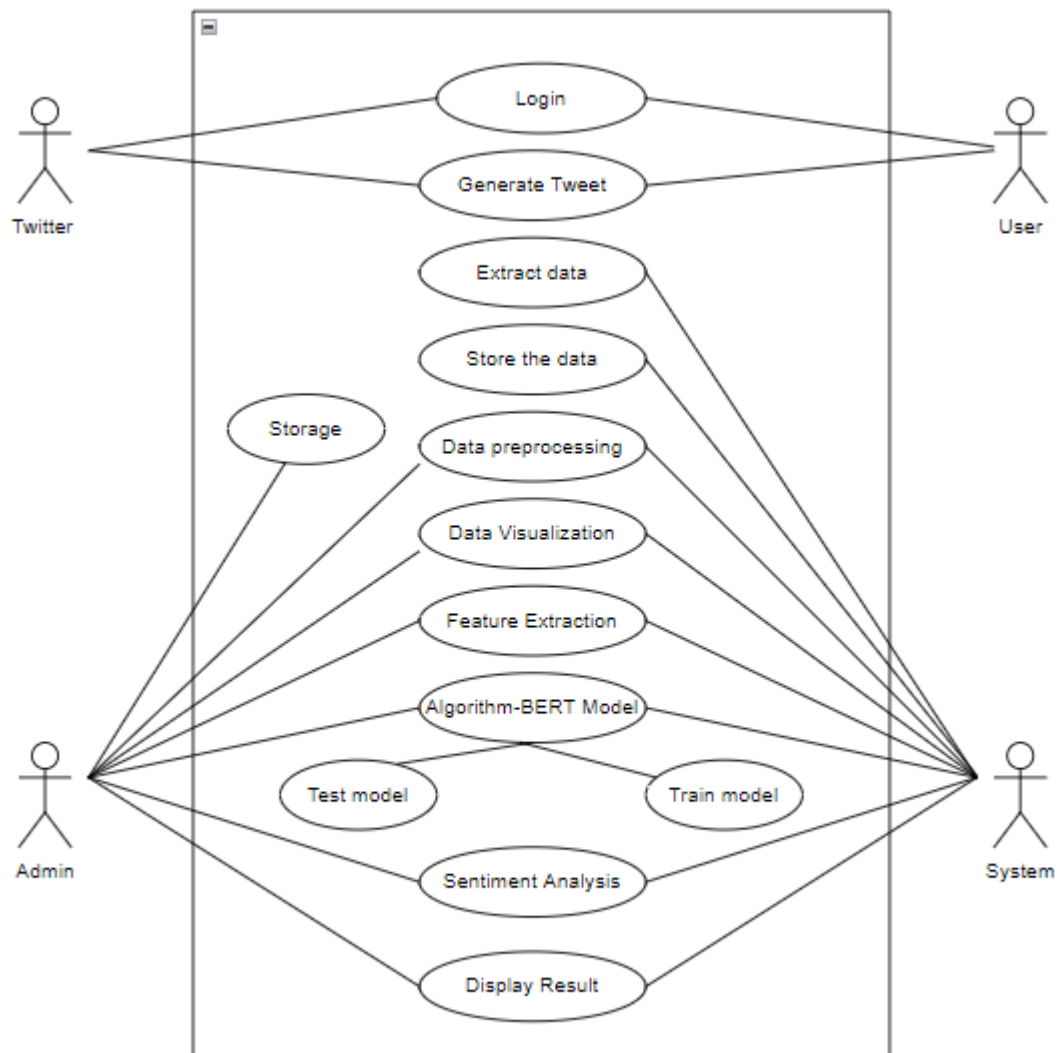
Tanisha Nazare(59)

Somu Sharma(50)

Ketaki Mankar(63)

Analysis Models

- **Use Case Diagram:**



The importance of using the use case diagram in our Project is that it would help us in capturing core functionalities of a system and visualize the interactions of various things called as actors with the use case. The use case in our Project is used for 4 actors including Twitter, User, Admin and the System itself. We are doing a sentiment analysis on the basis of news topics during COVID-19 using Twitter. The use case flow starts with first the user needs to login on the Twitter as well as it would generate a tweet then the system would collect and extract the data i.e. the tweets and would store the given data. After that data pre-

processing would take place where the raw data would be transformed to useful data. Then Data Visualization is done by admin to the system where graphical representation of data is done in the form of charts, graphs, maps, etc. Feature Extraction where raw data is transformed into numerical features to reduce the amount of redundant data. Then after applying the algorithm, we test and train the model and do the sentiment analysis on it and at last display the result.

Use Case tables:

- For User and Admin and for the functionality of Storing the data:

Use Case	Input
Primary Actor	User
Goal	Receive input from the user and pass onto the algorithm for detection.
Preconditions	Data should not be noisy.
Trigger	When user wants to detect the tweets
Scenario	To detect the tweets
Exceptions	-
Priority	Essential
When Available	Always
Frequency of Use	Regular
Channel to Actor	System
Secondary Actor	Admin
Channels to Secondary Actors	System

- For User and Admin and for the functionality of Data Visualization:

Use Case	Data Visualization
Primary Actor	Admin and System
Goal	Representation of tweets in the form of world maps,graphs,charts,etc.
PreConditions	Data should be accurate.
Trigger	When data is too large.
Scenario	When admin wants to visualize data from Twitter
Exceptions	Raw data of Tweets(Date,Time,UserName,Location)
Priority	Depends
When Available	Always
Frequency of Use	Regular
Channel to Actor	System
Secondary Actor	Admin
Channels to Secondary Actors	System

- For User and Admin and for the functionality of Data Preprocessing:

Use Case	Data PreProcessing
Primary Actor	User
Goal	Receive raw tweet for cleaning and to give ML model its

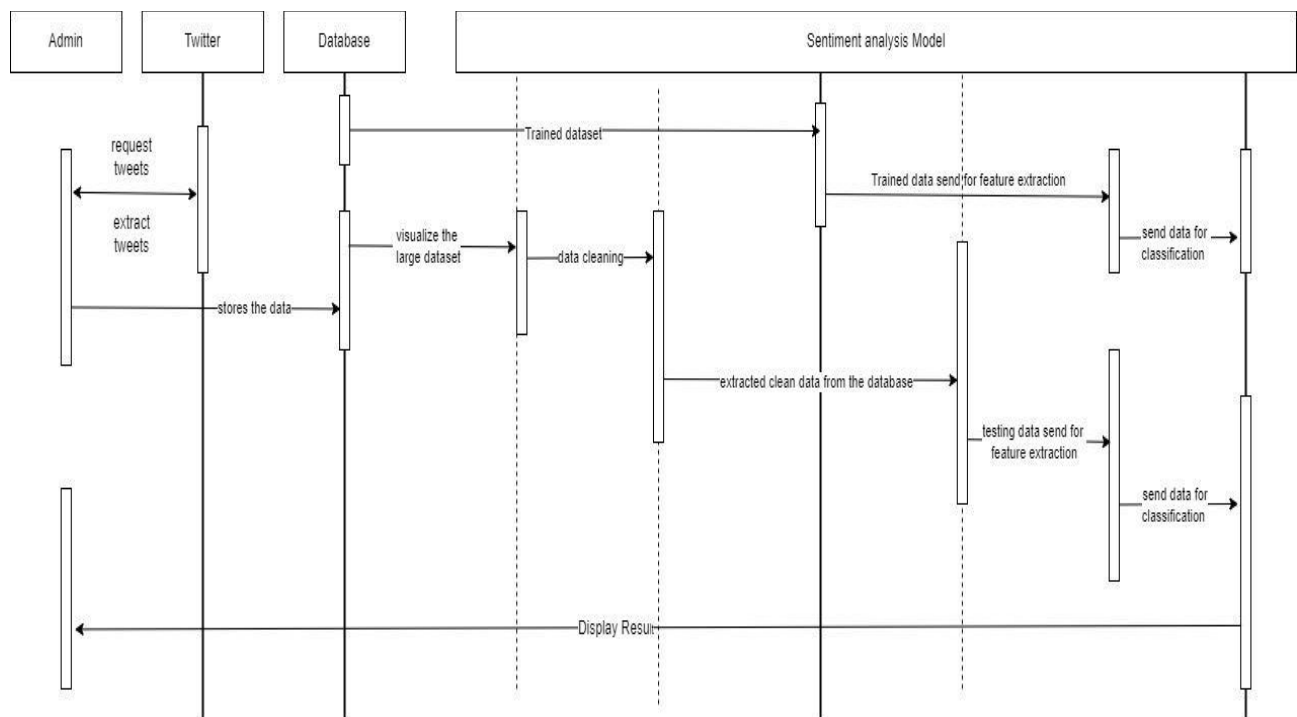
	appropriate data.
PreConditions	Data should be raw.
Trigger	When user wants to clean the raw data
Scenario	To detect the tweets
Exceptions	Tweets
Priority	Essential
When Available	Always
Frequency of Use	Regular
Channel to Actor	System
Secondary Actor	Admin
Channels to Secondary Actors	System

- For User and Admin and for the functionality of Sentiment Analysis:

Use Case	Sentiment Analysis
Primary Actor	Admin and System
Goal	Analysis of Tweets,its sentiments and emotions
PreConditions	Data should be accurate.
Trigger	When data is too large.
Scenario	When user wants to see the analysis of sentiments
Exceptions	Raw data doesn't work for ML model
Priority	Essential
When Available	Always

Frequency of Use	Regular
Channel to Actor	System
Secondary Actor	Admin
Channels to Secondary Actors	System

Sequence diagram:

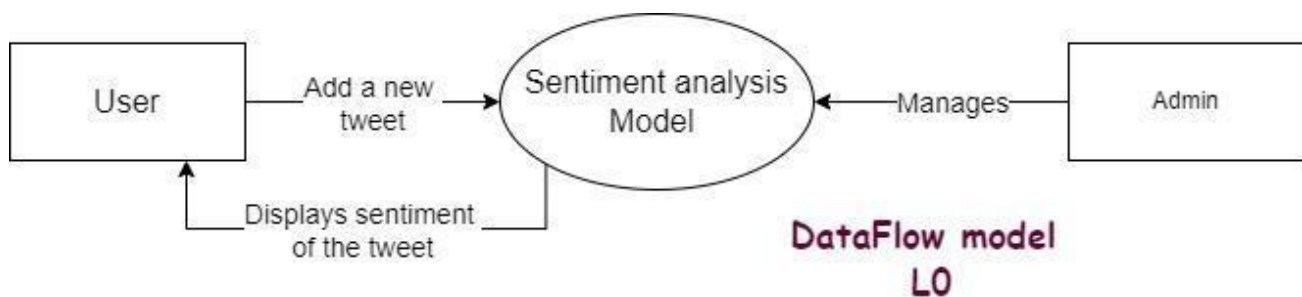


A sequence diagram is a type of interaction diagram because it describes how and in what order a group of objects works together. As can be seen at 1st we have a user who is going to interact with twitter and database. The user will extract data such as tweets and its information from twitter. that data is going to be stored in the database in the form of a dataset(.csv file). Then that data will be sent to the sentiment analysis model for analysis of

data and predicting the sentiment of that data. The model also visualizes the data stored in the dataset6 in the form of graphs, world map and word cloud.

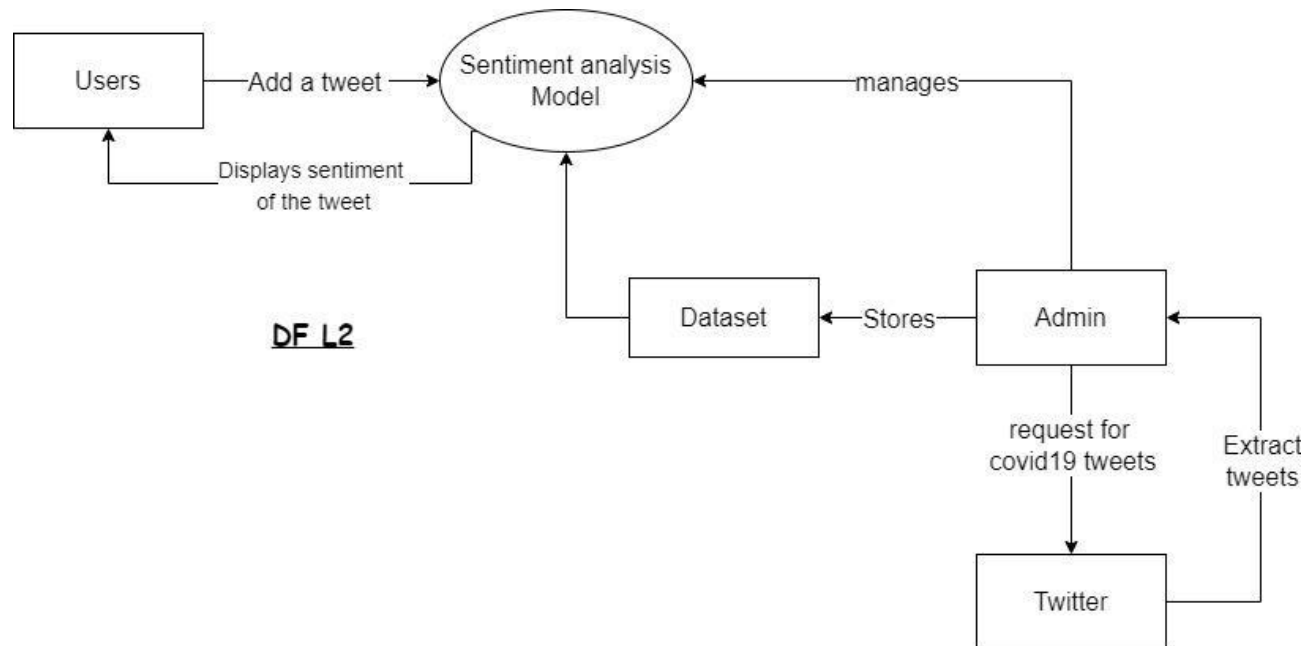
- **Data Flow Diagram:**

Level 0:



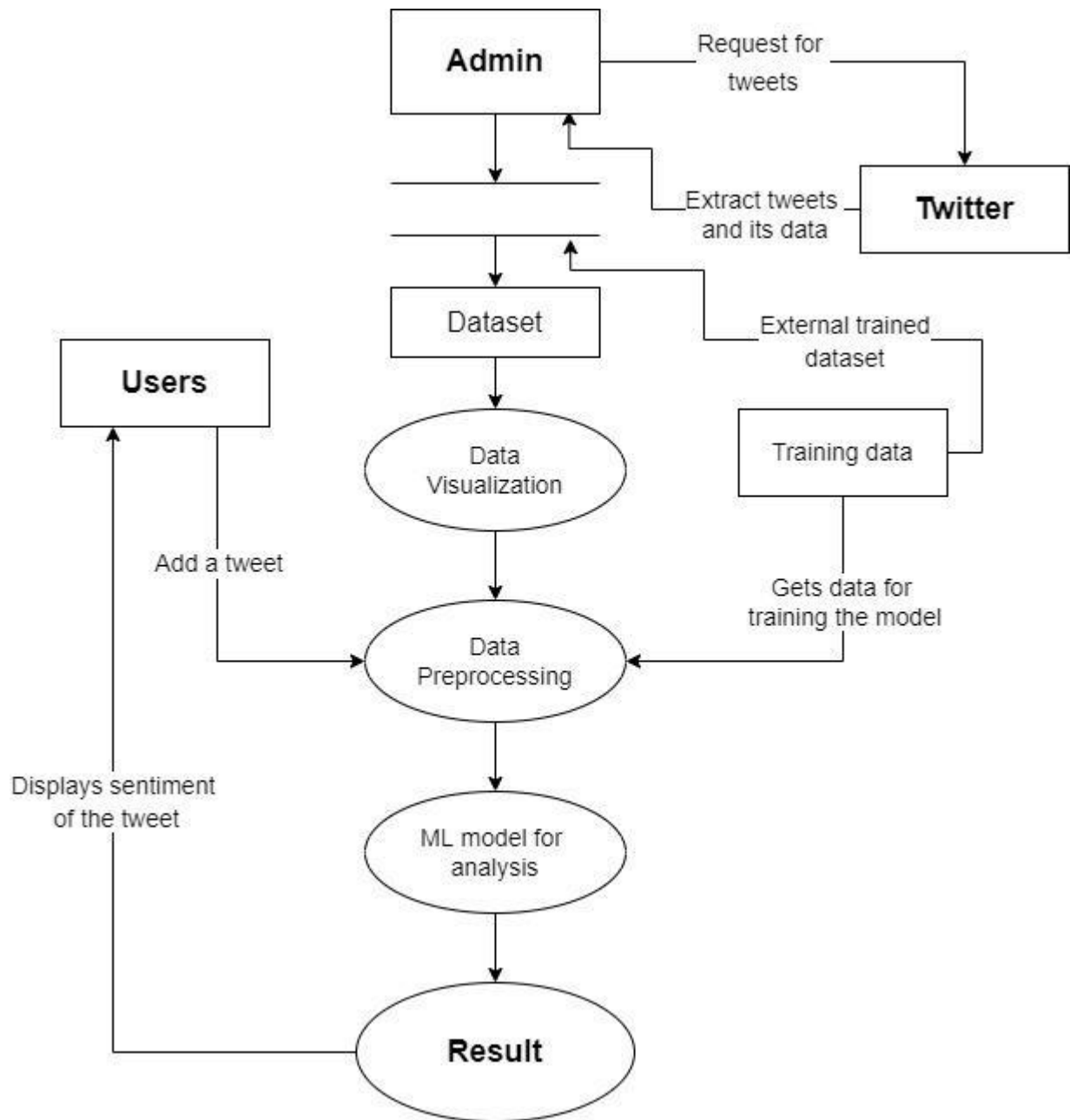
It is also known as context diagram. It basically shows the system as a single process with its relationship to external entities by incoming/outgoing arrows.

Level 1:



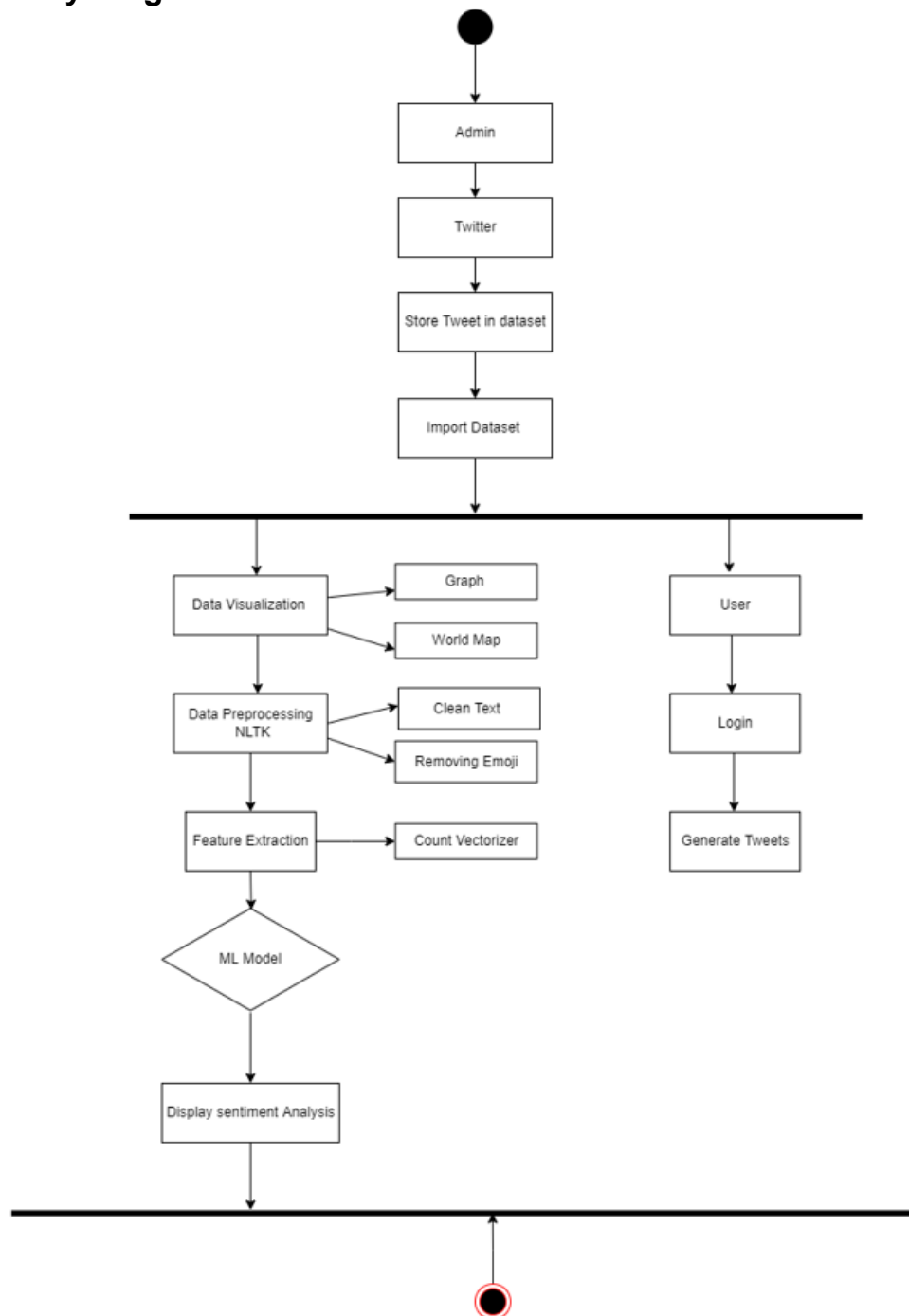
In level-1 DFD, the level-0 diagram is decomposed into multiple processes and we highlight the main functions and break the high level processes into subprocesses.

Level 2:



2-level DFD goes one step deeper into parts of 1-level DFD. It can be used to plan or record the specific/necessary detail about the system's functioning. the data-flows to and from the process on the level 1 diagram will all become data-flows across the system boundary on the level 2 diagram.

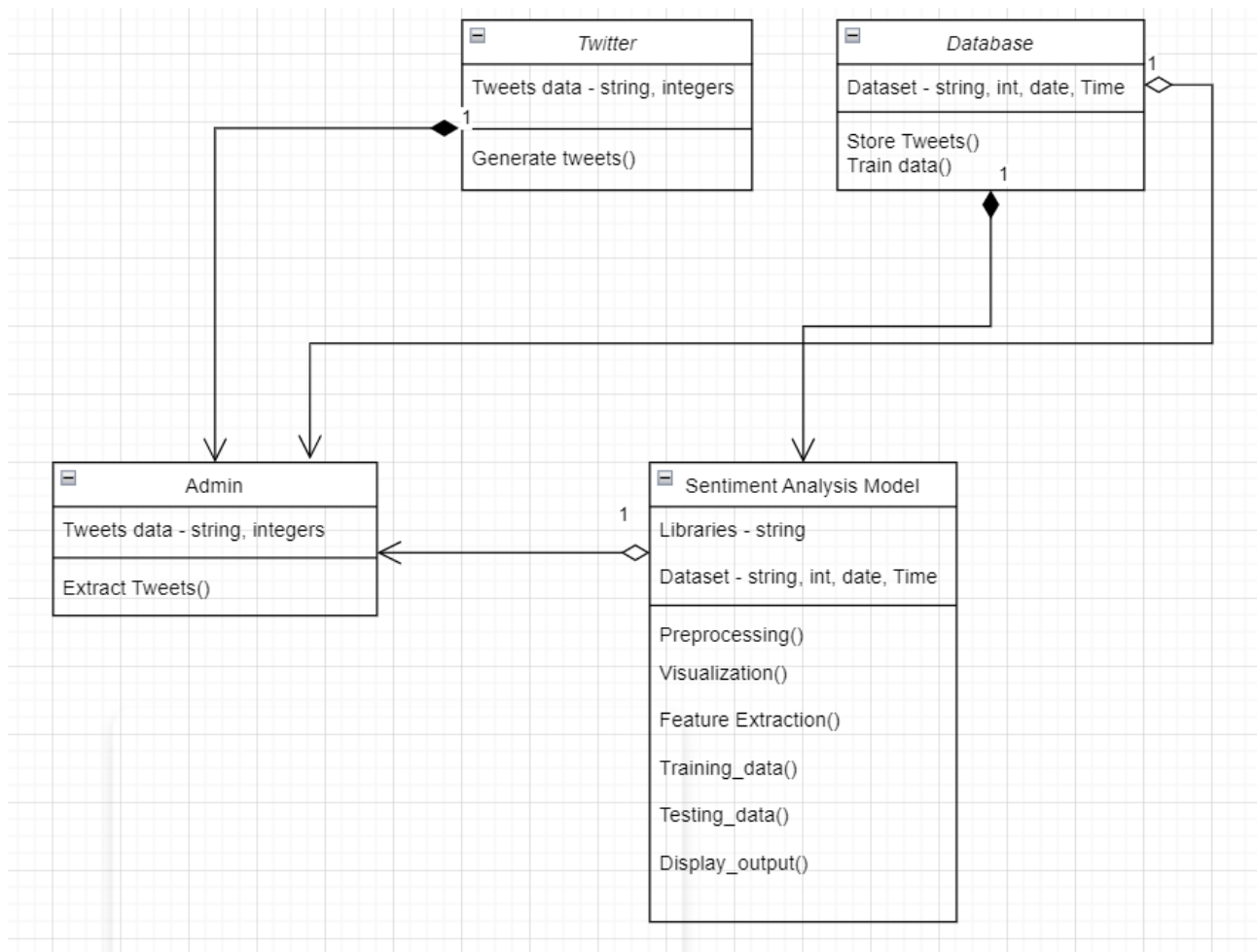
- **Activity Diagram:**



In the Unified Modeling Language, activity diagrams are intended to model both computational and organizational processes. Activity diagrams show the overall flow of control. Here activity in our Project is performed for 2 actors including the User and Admin. Using Twitter sentiment analysis is performed on the basis of news topics during COVID-

19. Here the user plays an important role in generating tweets. The user will login to twitter and will generate Tweets based on the generated tweets admin will perform the sentiment analysis. The flow starts with the admin collecting tweets in the dataset. Then Data Visualization is done by admin to the system where graphical representation of data is done in the form of charts, graphs, maps, etc. Feature Extraction where raw data is transformed into numerical features to reduce the amount of redundant data. By applying the ML algorithm model will be created and the result will display sentiment analysis.

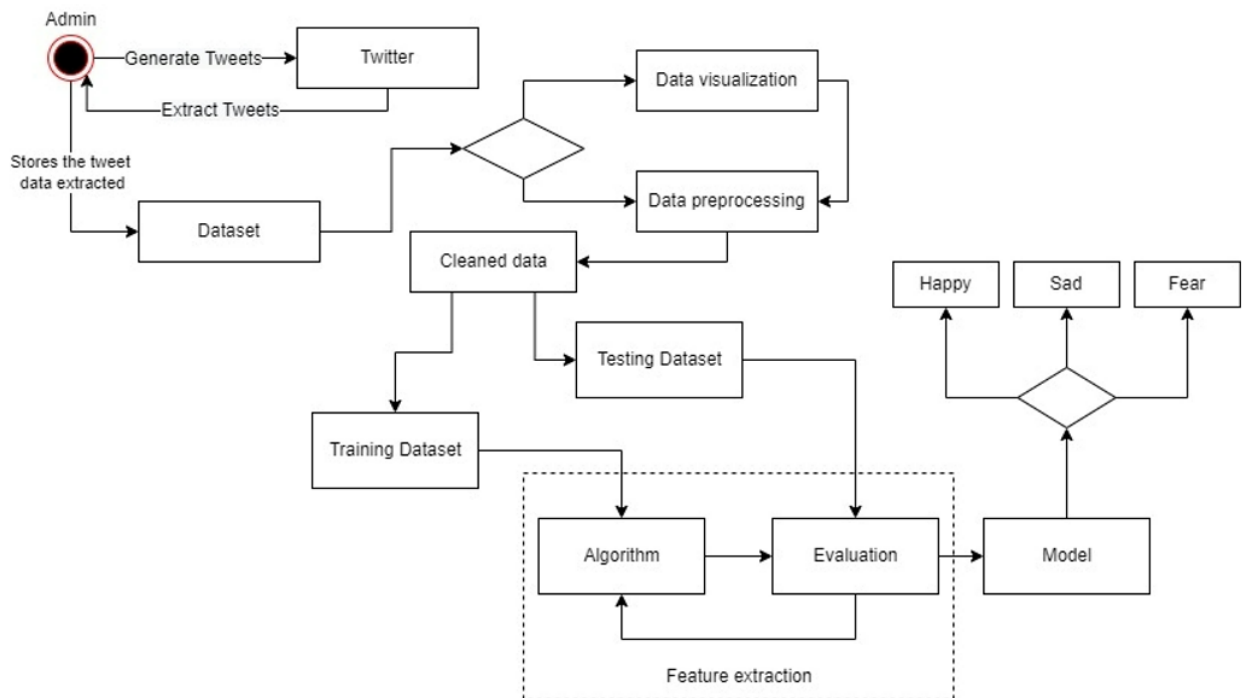
- **Class Diagram:**



We use a class diagram for modeling and displaying the relationships between the users and in our class diagram, it starts with Twitter from where we fetch the data and generate the tweets given by the user and also we fetch the data from an external data sources available on medias like kaggle,etc in the form of tweets. The admin will then extract the tweets and in our model with the help of the libraries having data type string and dataset having data type containing string, integer, date and time, the sentiment analysis model basically inherits the data from the database and from Twitter then we will do functions like preprocessing the

dataset and then visualizing the results in terms of graphs then feature extraction for getting the numerical features the we will train and test the dataset and would display the output.

- **State Diagram:**



A state diagram is a diagram used to describe the behavior of a system considering all the possible states of an object when an event occurs. State diagrams graphically represent finite state machines. They are only used to understand object behavior throughout the whole system. Initially the user will login twitter then the user will generate tweets, the tweets will get extracted into a dataset and hence will get stored. Data preprocessing is done on the Training dataset to remove null values. Data Visualization is done by admin to the system where graphical representation of data. Feature Extraction where raw data is transformed into numerical features. After applying sentiment analysis the analyzed sentiment will be displayed in the form of emotion like happy, Joy, Sad.