REAL TIME TOP TRENDING TWEET ANALYSIS

Author

Capt Nabeel Ahmed

Registration Number
00000241018

Supervisor

Maj Khawir Mehmood

Department of CSE, MCS

Submitted to the faculty of Department of Computer Software Engineering, Military College of Signals, National University of Sciences and Technology, in partial fulfillment for the requirements of B.E Degree in Computer Software Engineering, June 2021

REAL TIME TOP TRENDING TWEET ANALYSIS

By

Capt Nabeel Ahmed

Capt Munir Hussain

Capt Zeeshan Asif

Capt Hamza Qayyum

Supervisor

Maj Khawir Mehmood

Department of CSE, MCS

Submitted to the faculty of Department of Computer Software Engineering, Military College of Signals, National University of Sciences and Technology, in partial fulfillment for the requirements of B.E Degree in Computer Software Engineering, June 2021

CERTIFICATE OF CORRECTIONS & APPROVAL

Certified that work contained in this thesis titled *"Real Time Top Trending Tweet Analysis"* carried out by Capt Nabeel Ahmed, Capt Munir Hussain, Capt Zeeshan Asif, Capt Hamza Qayyum under the supervision of Maj Khawir Mehmood for partial fulfillment of Degree of Bachelors of Computer Software Engineering, in Military College of Signals, National University of Sciences and Technology, Islamabad during the academic year 2020-2021 is correct and approved. The material that has been used from other sources it has been properly acknowledged / referred.

Approved by

Col Asim Dilawar Bakhshi

Supervisor

Maj Khawir Mehmood

Department of CSE, MCS

Date: _____

ABSTRACT

REAL TIME TOP TRENDING TWEET ANALYSIS

RETTA is primarily a data science based project that include data analysis and data visualization of bunch of real time tweets (that will be scrapped from the twitter using a web scrapper) of the specific trend. It will comprise of a dashboard, on which there will be a button to scrape top trends on getting trends we can choose one among twenty trends and get analysis report. Users of this project will be some marketing agencies, decision making department of a company, government or general public. This project is web based application which will use web scrapper to scrape real time data; NLP model to analyze tweets and google charts to visualize result. This back-end is developed on REST API architecture.

DECLARATION

No portion of work presented in this thesis has been submitted in support of another award or qualification in either this institute or anywhere else.

Plagiarism Certificate (Turnitin Report)

This thesis has been checked for Plagiarism. Turnitin report endorsed by Supervisor is attached.

turnitin

Digital Receipt

This receipt acknowledges that Turnitin received your paper. Below you will find the receipt information regarding your submission.

The first page of your submissions is displayed below.

Submission author:	Nabeel Ahmed
Assignment title:	Plagiarism detection (June 2021) Part 2 (Moodle TT)
Submission title:	Real Time Top Trending Tweets Analysis (RETTA)
File name:	96617_Nabeel_Ahmed_Real_Time_Top_Trending_Tweets_Ana...
File size:	2.99M
Page count:	65
Word count:	6,458
Character count:	34,579
Submission date:	16-Jun-2021 03:25AM (UTC-0700)
Submission ID:	1607410714

REAL TIME TOP TRENDING TWEET ANALYSIS

Author
Capt Nabeel Ahmed

Registration Number
00000241018

Supervisor
Maj Khawir Mehmood
Department of CSE, MCS

Submitted to the faculty of Department of Computer Software Engineering, Military College of Signals, National University of Sciences and Technology, in partial fulfillment for the requirements of B.E Degree in Computer Software Engineering, June 2021

Copyright 2021 Turnitin. All rights reserved.

Real Time Top Trending Tweets Analysis (RETTA)

ORIGINALITY REPORT

8%	5%	0%	6%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Higher Education Commission Pakistan Student Paper	2%
2	manualzz.com Internet Source	1%
3	Submitted to Coventry University Student Paper	1%
4	Submitted to Australian Institute of Higher Education Student Paper	<1%
5	Submitted to Charotar University of Science And Technology Student Paper	<1%
6	Submitted to Noroff University College Student Paper	<1%
7	Submitted to Kingston University Student Paper	<1%
8	irigs.iiu.edu.pk:64447 Internet Source	<1%

9	lup.lub.lu.se Internet Source	<1 %
10	Submitted to University of Paisley Student Paper	<1 %
11	Submitted to Western Illinois University Student Paper	<1 %
12	Submitted to University of London External System Student Paper	<1 %
13	jurnal.ubd.ac.id Internet Source	<1 %
14	nl.wikipedia.org Internet Source	<1 %
15	Submitted to South Tyneside College, Tyne & Wear Student Paper	<1 %
16	www.cdsmodel.com Internet Source	<1 %
17	www.coursehero.com Internet Source	<1 %
18	123dok.com Internet Source	<1 %
19	scholar.sun.ac.za Internet Source	<1 %

20	www.slideshare.net Internet Source	<1 %
21	es.scribd.com Internet Source	<1 %
22	link.springer.com Internet Source	<1 %

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	Off		

Signature of Student

Capt Nabeel Ahmed

Registration Number

00000241018

Signature of Supervisor

Maj Khawir Mehmood

Department of CSE, MCS

Acknowledgements

I am thankful to my Creator Allah Subhana-Watala to have guided me throughout this work at every step and for every new thought which You setup in my mind to improve it. Indeed, I could have done nothing without Your priceless help and guidance. Whosoever helped me throughout the course of my thesis, whether my parents or any other individual was Your will, so indeed none be worthy of praise but You.

I am profusely thankful to my beloved parents who raised me when I was not capable of walking and continued to support me throughout in every department of my life.

I would also like to express special thanks to my supervisor Maj Khawir Mehmood for his help throughout my thesis and also for Software Quality Engineering course which he has taught me. I can safely say that I haven't learned any other engineering subject in such depth than the ones which he has taught.

I would also like to pay special thanks to Asst Prof Mobeena Shahzad for her tremendous support and cooperation. Each time I got stuck in something, she came up with the solution. Without her help I wouldn't have been able to complete my thesis. I appreciate her patience and guidance throughout the whole thesis.

I would also like to thank Maj Zeeshan Zulkifl for being on my thesis guidance and evaluation committee and express my special thanks to Prof Athar Mohsin Zaidi for his help. I am also thankful to Maj Wajahat Sultan and Asst Prof Naima for their support and cooperation.

Finally, I would like to express my gratitude to all the individuals who have rendered valuable assistance to my study.

Dedicated to my exceptional parents and adored siblings whose tremendous support and cooperation led me to this wonderful accomplishment

PREFACE

Thesis extant requirement, design, testing and implementation of the project ***"Real Time Top Trending Tweet Analysis"***. For convenience we have distributed it into five chapters.

First Chapter: Outlines introduction, objectives and scope of the project

Second Chapter: The goal of requirement analysis is to gather information and requirements from stakeholders and users in order to design software. The scope, objective, functional, and non-functional requirements of the system are all included in this chapter.

Third Chapter: Provides detailed design of the application of the system and the objects in the system. The development techniques and interfaces of our system are discussed in this chapter.

Fourth Chapter: It covers software requirement specifications, in which each module by unit testing explains its functionality, behavior, performance, and quality. It is concerned with the system's environment and deployment settings, and it comprises all information about the system's use and actual manifestation.

Fifth Chapter: The user manual, which provides all instructions on how to use the software, is found in this chapter.

Table of Contents

CERTIFICATE OF CORRECTIONS & APPROVAL ..... **iii**
DECLARATION ..... **v**
Plagiarism Certificate (Turnitin Report) ..... **vi**
Acknowledgements ..... **x**
Table of Contents ..... **xiii**
Table of Figures **1**
CHAPTER 1 ..... **2**
CHAPTER 1: INTRODUCTION **3**
1.1 Background of RETTA 3
1.2 Objectives of RETTA 3
1.3 Significance of RETTA 3
CHAPTER 2 ..... **4**
REQUIREMENT ANALYSIS **4**
CHAPTER 2: REQUIREMENT ANALYSIS **5**
2.1 Purpose 5
2.2 Definition, Abbreviations and Acronyms 5
2.3 Project Scope 6
3 Overall Description 6
3.1 RETTA Perspective 6
3.2 RETTA Functionalities 6
3.3 User Classes and Characteristics 7
3.3.1 Summary of User Class Data Analysis Agent 7
3.3.2 User / Visitor 7
3.3.3 Data Analysis Agent 7
3.4 Operating Environment 8
3.5 Limitations in Design and Implementation 8
3.6 Developing 9
3.7 User Documentation 9
3.8 Assumptions and Dependencies 9
4 External Requirements for Interface 9
4.1 UI or User Interface 9
4.2 Hardware Interfaces 10
4.3 Software Interfaces 10
5 Salient Features of RETTA 11
5.1 Sub RETTA Features 11
5.2 User can Input Choices 11

5.2.1 Normal Path: 11
5.2.2 Exceptional Path: Error messages will be shown or default values will be selected. 12
5.3 Registered User 12
5.3.1 Normal Path: *User can use system after successful login and registration.* 13
5.4 Scrapper will Scrap Twitter's Data 14
5.4.1 Normal Path: Scraper will scrape data 14
5.4.2 Exceptional Path: Scraping will be Unsuccessful 14
6 Non-Functional Requirements 15
6.1 Requirements for Performance 15
6.2 Safety and Security Requirements 15
7 Software Quality Attributes 15
7.1 Usability 15
8 UML Diagrams 15
CHAPTER 3 17
RETTA DESIGN 17
CHAPTER 3: RETTA DESIGN 18
8.1 Definitions, Abbreviations and Acronyms 18
8.2 Purpose 18
8.3 Project Scope 19
9 System Overview 19
9.1 RETTA Perspective 19
9.2 Description of System Architecture 19
9.3 Structure and Relationships 19
9.4 UML Class Diagram 19
9.5 UI Issues 20
9.6 Diagram of Activity 20
9.7 Diagram of Sequence 20
9.8 UI Design 20
9.9 Detailed Description of Components 21
10 System Architecture 21
10.1 Modules Overview 21
10.2 Description of Diagram 22
10.2.1 Users 22
10.2.2 Processing 23
10.2.3 Web Scrapper 23
10.2.4 Sentiment Analyzer 23
10.2.5 Data Visualizer 23
10.3 Block Diagram 23
10.3.1 Block Diagram Description 24

10.4 Architecture Diagram 24
10.5 Use-Case Diagram 25
10.6 Diagram of Class 26
10.7 Activity Diagram 28
10.7.1 User Management 28
10.7.1 Web Scrapper 29
10.7.2 Data Analyzer 30
10.8 Sequence Diagram (Dynamic View) 31
10.8.1 Login 31
10.8.2 Web Scrapper 32
10.8.3 Data Analyzer 32
10.9 User-Interface Design 34
10.9.1 Description of Diagram 35
CHAPTER 4 41
QUALITY ASSURANCE 41
CHAPTER 4: QUALITY ASSURANCE 42
11 Test Items 42
11.1 Features Required to be Tested 42
11.2 Test Approaches 43
11.3 Pass or Fail Measure 43
11.4 Standard for Deferral and Renewal Requirements 44
11.5 Test Deliverables 44
11.5.1 User Interface Testing 44
11.6 Risk & Eventualities 54
11.6.1 Risk Scheduling 54
11.6.2 Risk in Operation 54
11.6.3 Risk in Technical Domain 54
11.6.4 Risk associated with Program 54
CHAPTER 5 55
USER MANUAL 55
CHAPTER 5: USER MANUAL 56
12 Installation and Usage Guide 56
APPENDIX A 57
APPENDIX B 58
APPENDIX C 59
Project Proposal 59
REFERENCES 61
BIBLIOGRAPHY 62

Table of Figures

Figure 1 Abstract Diagram 19
Figure 2 Block Diagram 22
Figure 3 Architecture Diagram 37
Figure 4 Use Case Diagram 39
Figure 5 Class Diagram 46
Figure 6 User Activity Diagram 28
Figure 6.1 Web Scrapper Activity Diagram 50
Figure 6.2 Data Analyzer Activity Diagram 57
Figure 7 Login Sequence Diagram 58
Figure 7.1 Web Scrapper Sequence Diagram 66
Figure 7.2 Data Analyzer Sequence Diagram 67
Figure 8 User-Interface Diagram 68

CHAPTER 1

INTRODUCTION

CHAPTER 1: INTRODUCTION

In this project, we analyze and show the top trending tweets in a certain region by building models for categorizing "tweets" into positive, negative, and neutral sentiment. We create models for two classification tasks: a binary task that divides sentiment into positive and negative categories, and a three-way task that divides sentiment into positive, negative, and neutral categories.

1.1 Background of RETTA

Previously if we want to see views of general public regarding a specific trend we will go to twitter and read tweets and assume some pattern in their tweets that are not accurate as sentiment analysis of bulk of tweets is not humanly possible. For handle this sort of issue in RETTA we are using modern data analyses and visualization techniques.

1.2 Objectives of RETTA

RETTA project is primarily be focusing on the automation using data scrapping, data analyses and data visualization techniques to scrape twitter. The analyzer will help the marketing agencies, customer services department, brand monitoring department, social media monitoring department and even celebrities or military to making their decisions and can check masses reviews on the basis of sentiment analysis of real time top trending tweets.

1.3 Significance of RETTA

- Scrape tweeter top twenty trends
- Scrape tweets regarding specific trends
- Data cleaning and prepossessing of scrapped data
- Data analysis of pre-possessed data
- Data Visualization on analysis result

CHAPTER 2

REQUIREMENT ANALYSIS

CHAPTER 2: REQUIREMENT ANALYSIS

2.1 Purpose

RETTA will be focusing on the sentiment analysis of textual data which will mainly tweets. It will visualize that analysis in the form of three chats which are pie chart, bar chart and line chart. It will also help users to resister themselves and use this app to know about opinion of masses.

2.2 Definition, Abbreviations and Acronyms

RETTA	Real Time Top Trending Tweets Analysis
CMS	Content Management System http://cms.nust.edu.pk/
NLP	Natural Language Processing
bs	Beautiful Soap
CSE Dept	Computer Software Engineering Department
HTML	Hypertext Markup Language
HTTP	Hypertext Transfer Protocol
HTTPS	Hypertext Transfer Protocol Secure
JWT	JSON Web Token
LMS	Learning Management System https://lms.nust.edu.pk
MCS	Military College of Signals
NUST	National University of Sciences and Technology
NLTK	Natural Language Tool Kit

2.3 Project Scope

RETTA (sentiment analyzer) will help the marketing agencies, customer services department, brand monitoring department, social media monitoring department and even celebrities or military to making their decisions and can check masses reviews on the basis of sentiment analysis of real time top trending tweets, given that previously partially observed, by generating tweets data statically and do sentiment analysis and data visualization of that data by hard core method using human intelligence and statistic techniques. It must be noted that this will be able to analyses and visualize real time top trending tweets of a selected country. As it is legal to scrape, analyses and visualize tweeters data available. Major implementation of the first version will be able to scrape data of the real time top trending tweets of a selected country, data analysis of these tweets to analyze sentiments using data science techniques and visualize these tweets on the basis of sentiments in the form of charts using data visualization techniques.

3 Overall Description

3.1 RETTA Perspective

This RETTA can be useful in following perspective:

- Easily accessible to all intended users as manually analysis is costly as it may include analyses using statics techniques and data gathering.
- It can be helpful for maintaining histories in the database.
- Improvements in RETTA as per feedback received from public.
- Addition of new features

3.2 RETTA Functionalities

The main Functional Requirements of RETTA are highlighted below:

- Scraping (tweeters data) top trending tweets (depends on user).
- Have option to select number of tweets of each trend to scrape with the specific country.
- Analyze these tweets on the basis of sentiments classification (Positive, Negative or Neutral) using data science technique called NLP (natural language processing).
- Data visualization of resulted insights in the form of graph or charts using chart.js.

- Have an option to save history form users, to maintain this we will use MongoDB as a No-SQL database.
- User interface will be managed and maintained using JavaScript and Ajax.
- Only registered users are able to access this web application

Following are the non-functional requirements:

- Easy to understand
- System must be able to display data with all desired actions
- Password protected

3.3 User Classes and Characteristics

3.3.1 Summary of User Class Data Analysis Agent

The following section describes about the of user and data analysis agent of the RETTA. There are explanations of the user followed by the interactions the analysis agent shall be able to make with the software.

3.3.2 User / Visitor

The user in this is basically the main and only user who can use RETTA. User must be a registered user and can use application only after authentication. Firstly, user inputs his choices on which data scraping will be take place. Then scraping data will be used by NLP agent (sentiment analyzer) and finally result will be shown in the form of tweets, charts (that show data classification) and reviews.

3.3.3 Data Analysis Agent

We will use NLP (a data science technique use for natural language processing) for sentiment data analysis. After that our data will be classified in to three groups on the basis of words.

World or Expression	Polarity
• e.g. Great, excellent, good, appreciable or :-) etc	• Positive
• e.g. Worst, bad, injurious or :-(etc	• Negative
• e.g. Doesn't matter, so so, or :-| etc	• Neutral

This analyses will be depending upon data set; more accurate data-set will lead us to more accurate data analyses.

3.4 Operating Environment

This is a web based application. The software and languages used are mentioned below:

- Visual Studio Code
- Python Language
- Chrome Web Driver
- Ubuntu / Windows
- Adobe Photoshop / Illustrator

The system of RETTA shall run on the computer system with following specifications

- Pentium 4 or higher CPU
- At least 2 GB RAM
- At least 1 GB free disk space
- Windows 10

3.5 Limitations in Design and Implementation

The goal of this project is to create an application that can scrape and analyze tweet's sentiment and visualize the findings in form of charts and graphs.

- Login page
- Basic event handlers
- Multiple event addition side by side

3.6 Developing

The most important module in this project is Data science based sentiment analysis for which we have used preprocessing and data cleaning approaches with NLTK (Natural language toolkit).

Key Features

Can support any kind of system or mobile phone.

Result from NLTK (Natural language toolkit) will be shown using pie chart.

3.7 User Documentation

Final release will be a web application that also deployed on AWS. User documentation that would be delivered along with the final product. User manual with instructions, pictures and text for understanding how to operate RETTA.

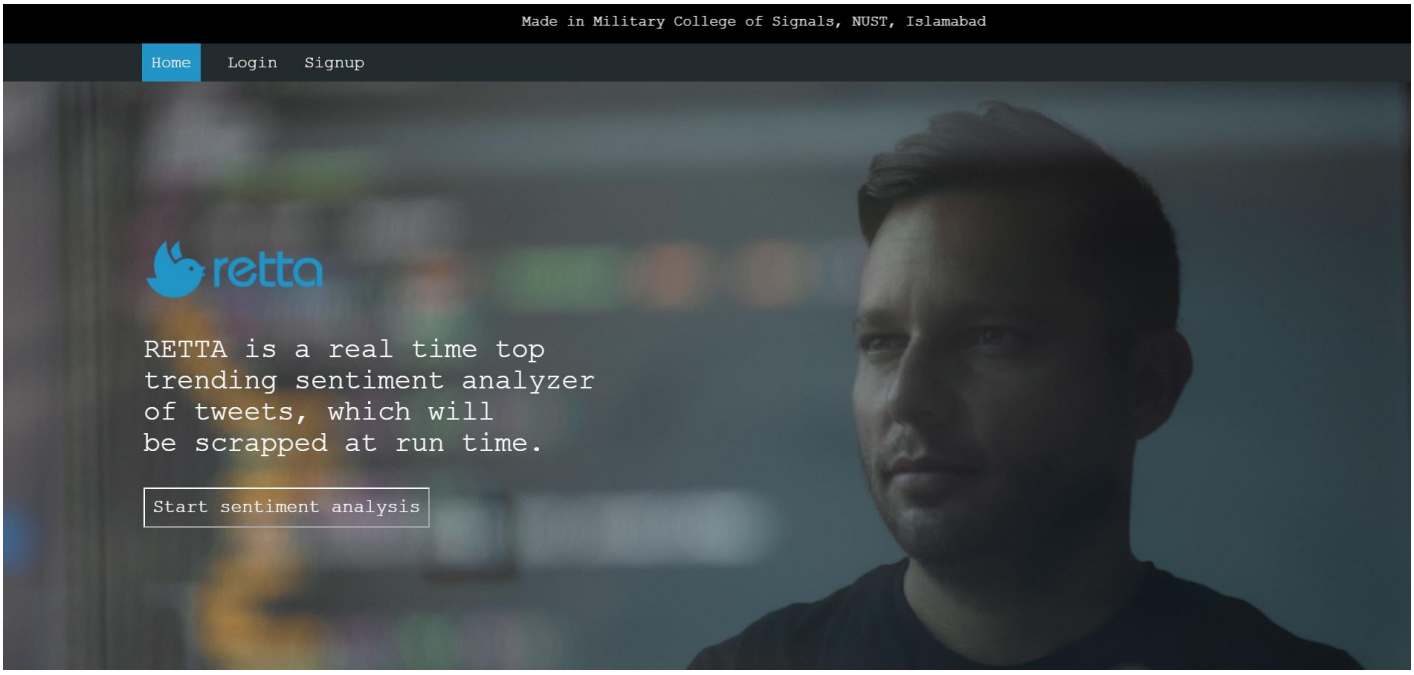
3.8 Assumptions and Dependencies

- Equipment in simulator will be taken as actual signals equipment operated in the field
- It is a standalone Web app so it can operate in any independent environment
- It does require Wi-Fi or internet access

4 External Requirements for Interface

4.1 UI or User Interface

UI available are the main front panel after login which can be handled after taking subsequent actions. Screenshots are shown below.



4.2 Hardware Interfaces

There will be an interface drawn for each panel of equipment designed in coral draw. User of RETTA will first login to the system and then choose a trend and click on the button to analyze trends and response will be shown in the form of a chart. Computer system hardware interfaces are as follows:

- System shall have keyboard/mouse input
- System shall have a monitor

4.3 Software Interfaces

Following software will be used in this project

- Visual studio code
- MongoDB
- Adobe Photoshop / Illustrator
- Selenium
- Beautiful Soap
- NLTK
- Chrome Driver

5 Salient Features of RETTA

This segment describes salient features of RETTA.

1. User can input their choices.
2. Only registered users can use this system.
3. Scraper will scrape twitters data.
4. Convert tweets into sentiment analysis charts.
5. Show these results in the form of charts.
6. Maintain user's history.

5.1 Sub RETTA Features

- After authentication the users are required to select a trend from a list if trends and then wait for almost 2 minutes, in that time tweets will be scrapped from the twitter and result will be shown in the form of pie-chart.
- Since it is Web based application it does depend on twitter and external communication channel like Wi-Fi, because it will use scrapper.

5.2 User can Input Choices

Use Case Requirement:

Use Case Paths
• Normal:
  User can input choices.
• Exceptional:
  If user didn't give proper values, default values will be selected.

5.2.1 Normal Path:
Externals

• No external resources
Preconditions • Default values will be selected.
Interactions • Input values will send to scraper.
Post conditions • Data visualization of the scraped data result.
5.2.2 Exceptional Path: Error messages will be shown or default values will be selected.
Externals • No external resources required.
Preconditions • Server side bug occurs so that server crash error **.**
Interactions An error message is sent to the system**.**
Post conditions • The error message is shown on the screen.

5.3 Registered User

Use Case Requirement: For using this system user must be a registered user.

Use Case Paths
• Normal:
  User can use this system after successful login or registration.
• Exceptional
  User can't able to use this system.

5.3.1 **Normal Path:** ***User can use system after successful login and registration.***
Externals • No external resources required.
Preconditions • User can use RETTA after that he can use system.
Post conditions • The system start generating the virtual map.

5.4 Scrapper will Scrap Twitter's Data

Use Case Requirement: Data scraping of tweeters data on the basis of top trending tweets.

Use Case Paths
• Normal:
  Scraper will scrape data successfully.
• Exceptional:
  Scraping will be Unsuccessful

5.4.1 Normal Path: Scraper will scrape data
Externals • No external resource required.
Preconditions • Registered user can only use this system.
Interactions • Scraped data will send to the NLP agent (sentiment analyzer).
Post conditions Result of the sentiment analyses of the tweets will be displayed in the form of charts.
5.4.2 Exceptional Path: Scraping will be Unsuccessful
Externals • No external resources required.
Preconditions • Non registered users are not able to use this system.
Interactions • Error messages will be displayed.
Post conditions • Scraping will be unsuccessful.

6 Non-Functional Requirements

6.1 Requirements for Performance

- Must have PC or mobile phone available.
- Chrome browser is a must.

6.2 Safety and Security Requirements

There is a not any specific security requirement as this is a web application. And users can be any one. For application security we are using JWT based authentication though that any unauthorized user will not have any accesses to our Restful API.

7 Software Quality Attributes

- It is easy to use (operability)
- It is easy to understand
- It is maintainable. (maintainability)
- It is secure. (security)
- It is compatible with any of Window/Linux/iOS (compatibility)

7.1 Usability

Usability is an important criterion in the development of RETTA. When a user selects a trend from a list after authentication and then scrapper will scrape tweets from the twitter. The GUI for data visualization is as easy as it is using different kind of charts.

8 UML Diagrams

- Appendix A: Basic Flow Chart / Diagram is attached.
- Appendix B: Use-Case Diagram is attached.

User Case ID	1
User Case Name	Analysis Report
Actors	Users, Twitter
Created By	Nabeel Ahmed
Date Created	12-06-2021
Description	Actor open application and select a trend and then press a button to analyze trend
Preconditions	The actor has to open RETTA on chrome browser
Post conditions	If the use case is successful the actor will able to see his or her analyses result in the form of pie chart.
Normal Flow	The use case start after user open RETTA on browser

CHAPTER 3

RETTA DESIGN

CHAPTER 3: RETTA DESIGN

This design document encapsulates all functional requirements and illustrates how they conceptually interact. The low-level design also demonstrates how these needs have been implemented. The document's purpose is to provide details about the design and the design process to stakeholders. This document will assist the developer in implementing the functionality and making it easier to comprehend.

8.1 Definitions, Abbreviations and Acronyms

Authentication	Authenticate all API using JWT token
Web Scrapper	Scrape top trends and tweets from twitter using chrome driver
Sentiment Analyzer	Analyze tweets text and tell weather tweets have positive, negative or neutral sentiments

8.2 Purpose

The goal of this Design Document is to offer a detailed explanation of the system's design so that software development may begin with a clear knowledge of what needs to be developed and how it should be developed. This Design Document contains the necessary information to create a detailed description of the software and system to be constructed. The goal of this article is to offer a design view of the RETTA as well as a full explanation of it. It will describe the system's goal and characteristics, as well as the system's interfaces, what the system will perform, its complete process in detail, the limitations under which it must work, and how the system will react to inputs and outputs. This paper is primarily meant for the system's stakeholders and developers.

8.3 Project Scope

The Scope of project is firstly to improve some functions in RETTA as per feedback received from some initial users and did some improvements in UI as well as in functionalities. As in previous version there is only one chart to show sentiment analysis but know there are three charts. Also one slider is including in the application.

9 System Overview

9.1 RETTA Perspective

This design document outlines all of our functional requirements and shows how they abstractly connect to one another. The low-level design also shows how we've been implementing and how we'll be implementing all of these requirements. For the time being, this low-level design does not address any non-functional criteria that our system has, as stated in the SRS Document.

9.2 Description of System Architecture

The overall architecture of the system is explained in this part, as well as the introduction of numerous components and subsystems. It is mainly supported by System Architecture diagram which shows an insider’s perspective of the system by describing the high level software components that perform the major functions to make the system operational.

9.3 Structure and Relationships

This section ponders upon the interrelationships and dependencies among various components. It is mainly described by a diagram which is further augmented by explanatory text. UML Class diagram also helps us understanding the system structure.

9.4 UML Class Diagram

UML Class diagram further manifests the description of low level components of the software that include data storage and state details, thus making the system adequately comprehensible.

9.5 UI Issues

The key concepts of the product's user interface are presented in this section. The part is explained by an overall graphic that is supplemented with explanatory text, but does not go into technical specifics. Moreover, Activity diagrams, Sequence diagrams, and UI Design diagrams also elaborate the User Interface issues in a more intelligible manner.

9.6 Diagram of Activity

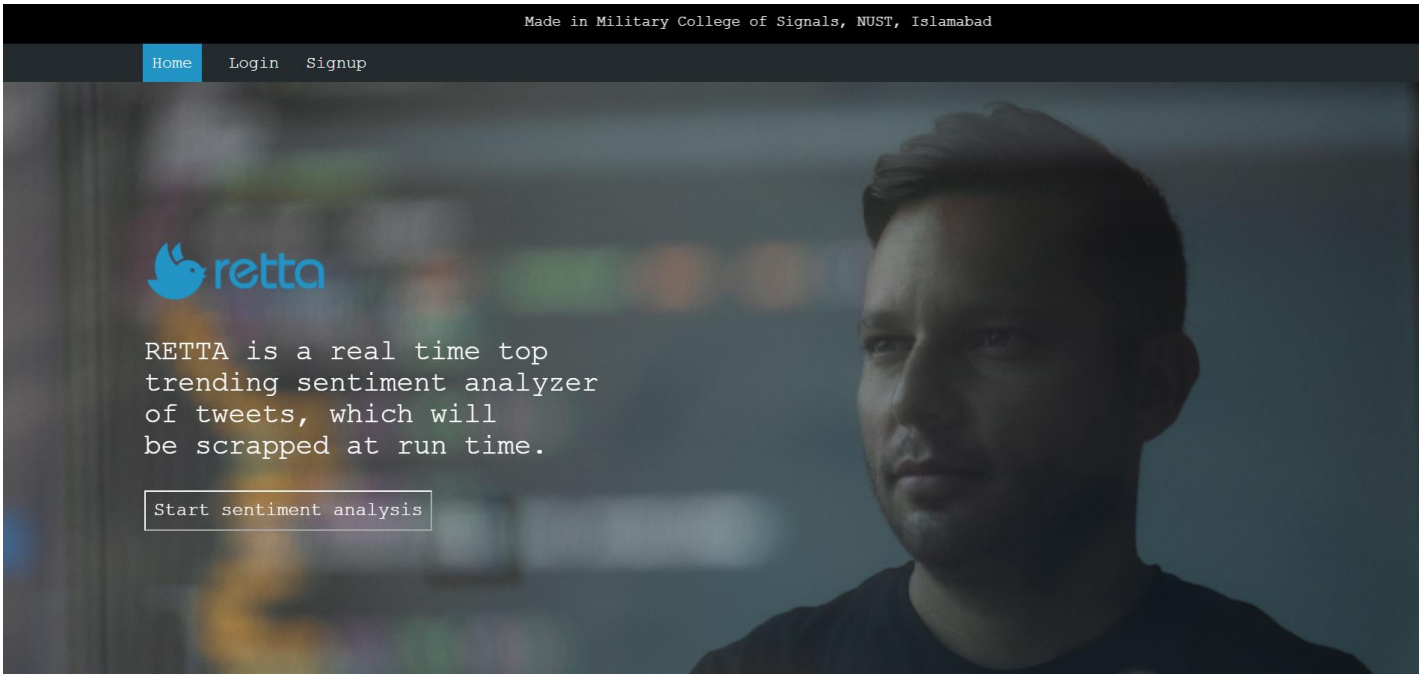
Activity Diagrams use a workflow-based method to represent the system's overall operation. They're a great way to illustrate how different stages are engaged in significant tasks inside a system that utilizes a flow chart design without having to go into the technical intricacies.

9.7 Diagram of Sequence

Sequence diagrams depict how various elements are engaged in the completion of a system's operation. They have a distinctive structure that allows the reader to observe how many objects are employed in relation to how long they take to complete a system need.

9.8 UI Design

Few images of GUI are present in this section that prototype the way a user shall be communicating with the system to get the desired results.



9.9 Detailed Description of Components

This section contains detailed description of all the major components of the system in a structured pattern (table), comprising of 10 x rows. The pattern (table) maintains symmetry in the document structure; and therefore it is followed for each of the components. Each part/row of the table is identified by a label, explaining the purpose of each point. The description of each point vis-à-vis the component being discussed, ponders upon the detailed account of it in the system.

10 System Architecture

In this section, the overall architecture of the system is discussed, including the introduction of various components and subsystems. It is mainly supported by System Architecture diagram which shows an insider's perspective of the system by describing the high level software components that perform the major functions to make the system operational.

10.1 Modules Overview

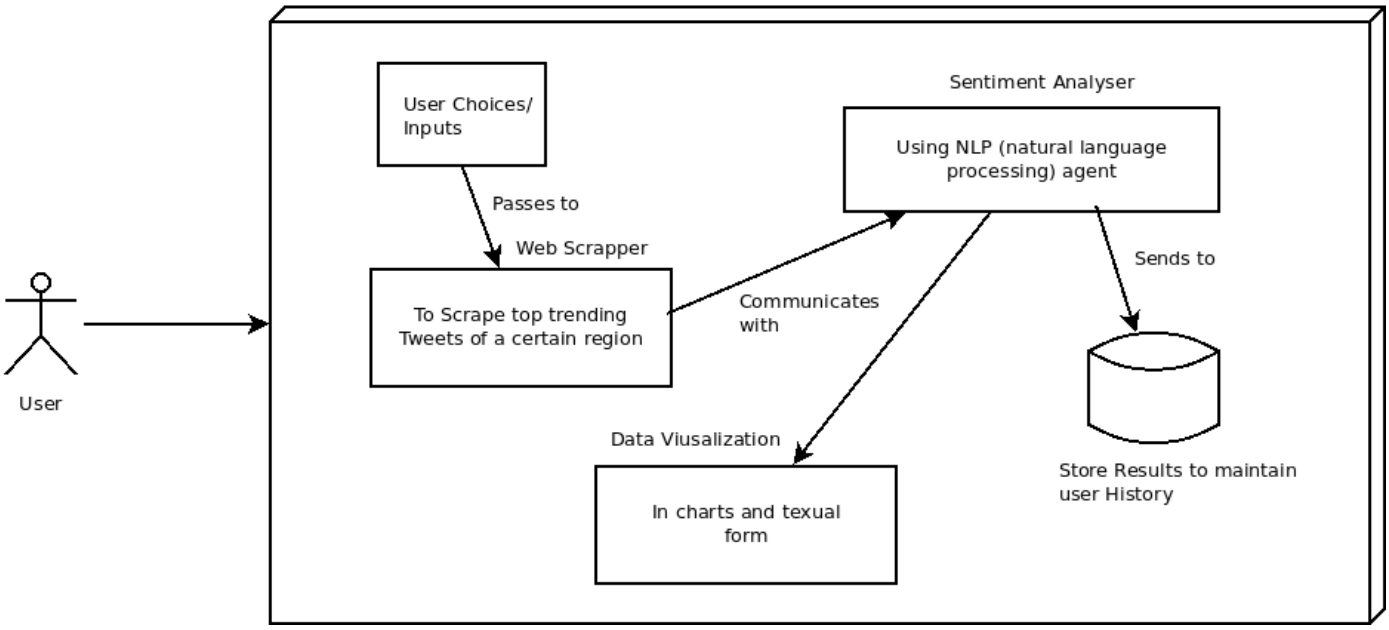


Figure 1 - Abstract Diagram

10.2 Description of Diagram

The system will be built around three primary modules: "Users," "Web scrapper," "NLP agent," and "Data visualizer." It will also have submodules, as depicted in the accompanying abstract diagram. The abstract picture depicts the entire system, from how people interact with it to how data is processed in databases. The Abstract diagram's sub modules are described in further detail below.

10.2.1 Users

Users of the RETTA will access the back-end System to login and then choose inputs/ choices according to his requirement. The user in this is basically the main and only user who can use RETTA. User must be a registered user and can use application only after authentication. Firstly, user inputs his choices on which data scraping will be take place. Then scraping data will be used by NLP agent (sentiment analyzer) and finally result will be shown in the form of tweets, charts (that show data classification) and reviews. User may be some marketing agencies, decision making department of a company, government or general public.

10.2.2 Processing

Back-end system will provide a platform to the user to interact and consume the services of the RETTA in an effective and more manageable manner.

10.2.3 Web Scrapper

The Scrapper will scrape real time top trending tweets from tweeter and store it in a data frame for further data cleaning and prepossessing in an efficient manner.

10.2.4 Sentiment Analyzer

The analyzer will analyze these data frame that we create as a resultant of scrapper and produce result that can visualize further by the visualizer.

10.2.5 Data Visualizer

This will visualize data (the resultant of analyzer) in the form of charts, text and shapes and store in a database to maintain user history.

10.3 Block Diagram

The principle parts of the RETTA are shown in the figure and their relation is defined by the connection between them. The details of all the modules is also given below:

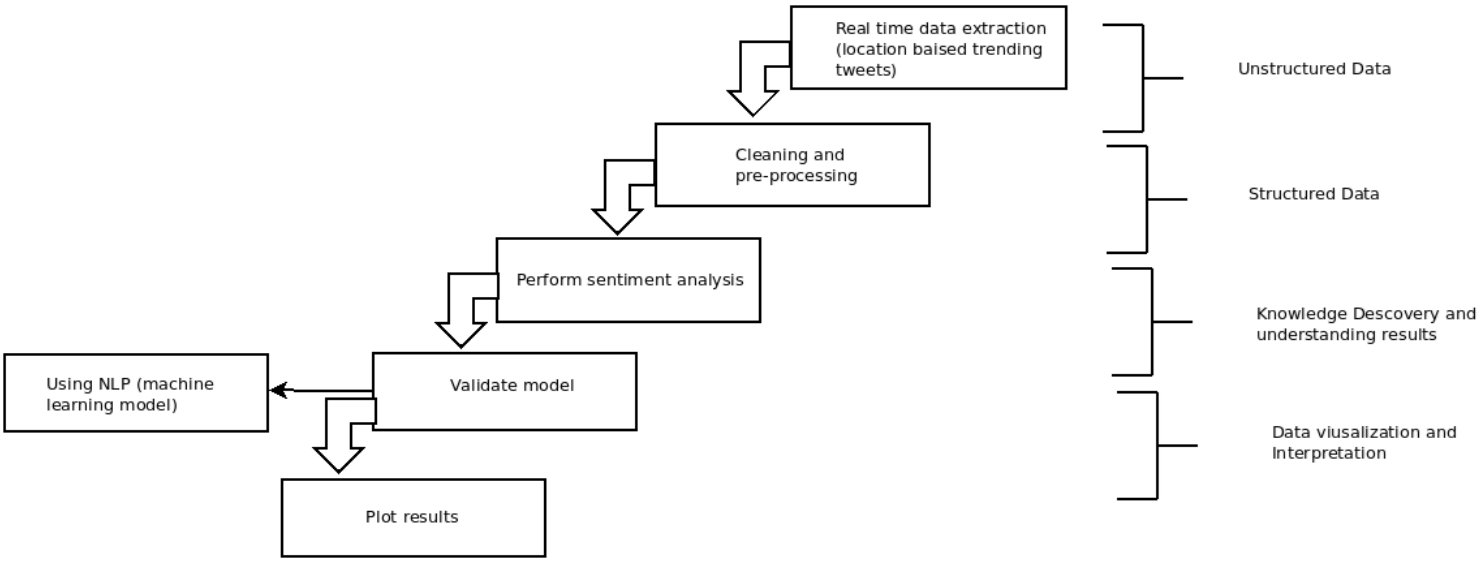


Figure 2 – Block Diagram

10.3.1 Block Diagram Description

- RETTA will only scrape top trending tweets.
- User can input his choices.
- It will only be used for sentiment analysis of data.
- Results from NLP agent will be visualize.
- User history will be stored, so that user can access it later.
- There can be time delay and less fluency.

10.4 Architecture Diagram

The main purpose of the project is to develop a RETTA is to help decision making process in different agencies on the basis of public views. The user can use the system to see the reaction of public towards various trends. It will offer the users to choose his location trends on that basis scraping and analysis occurs. The system is being built to automate the entire process, which is presently manual.

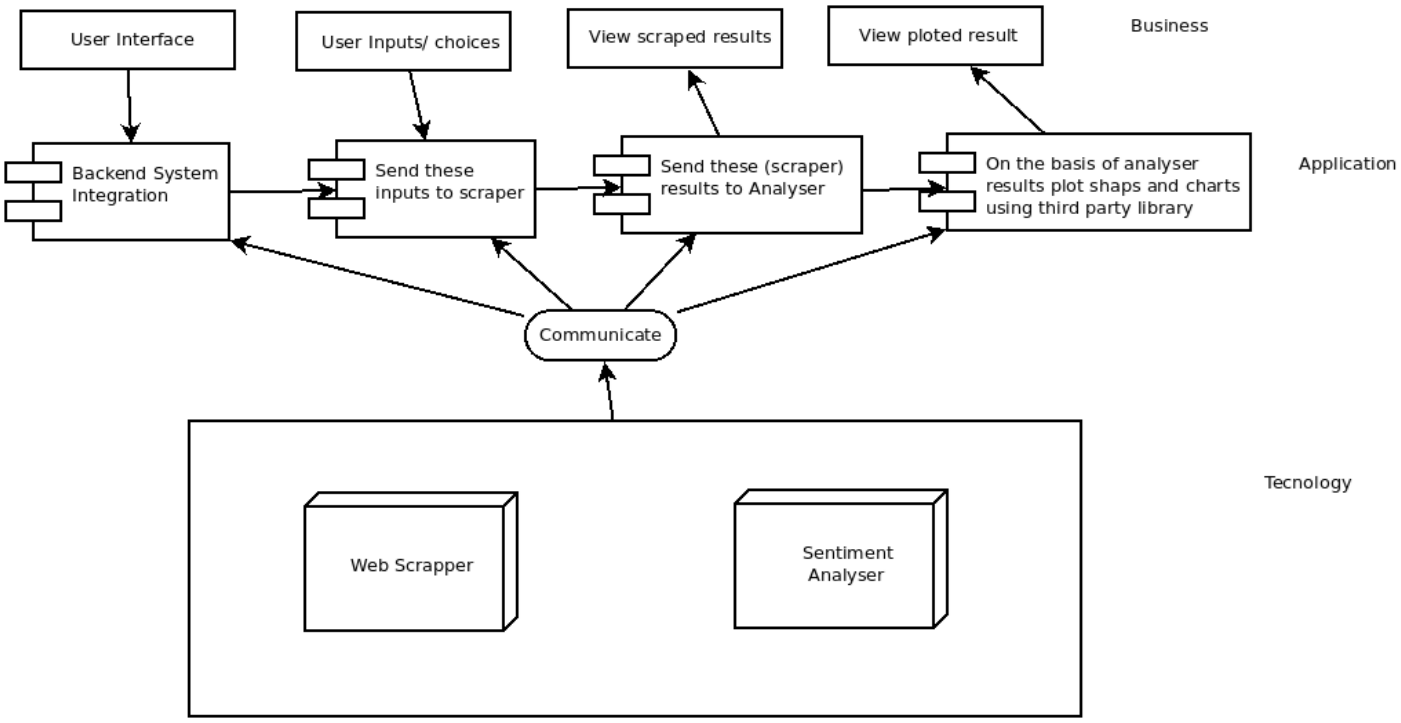


Figure 3 – Architecture Diagram

10.5 Use-Case Diagram

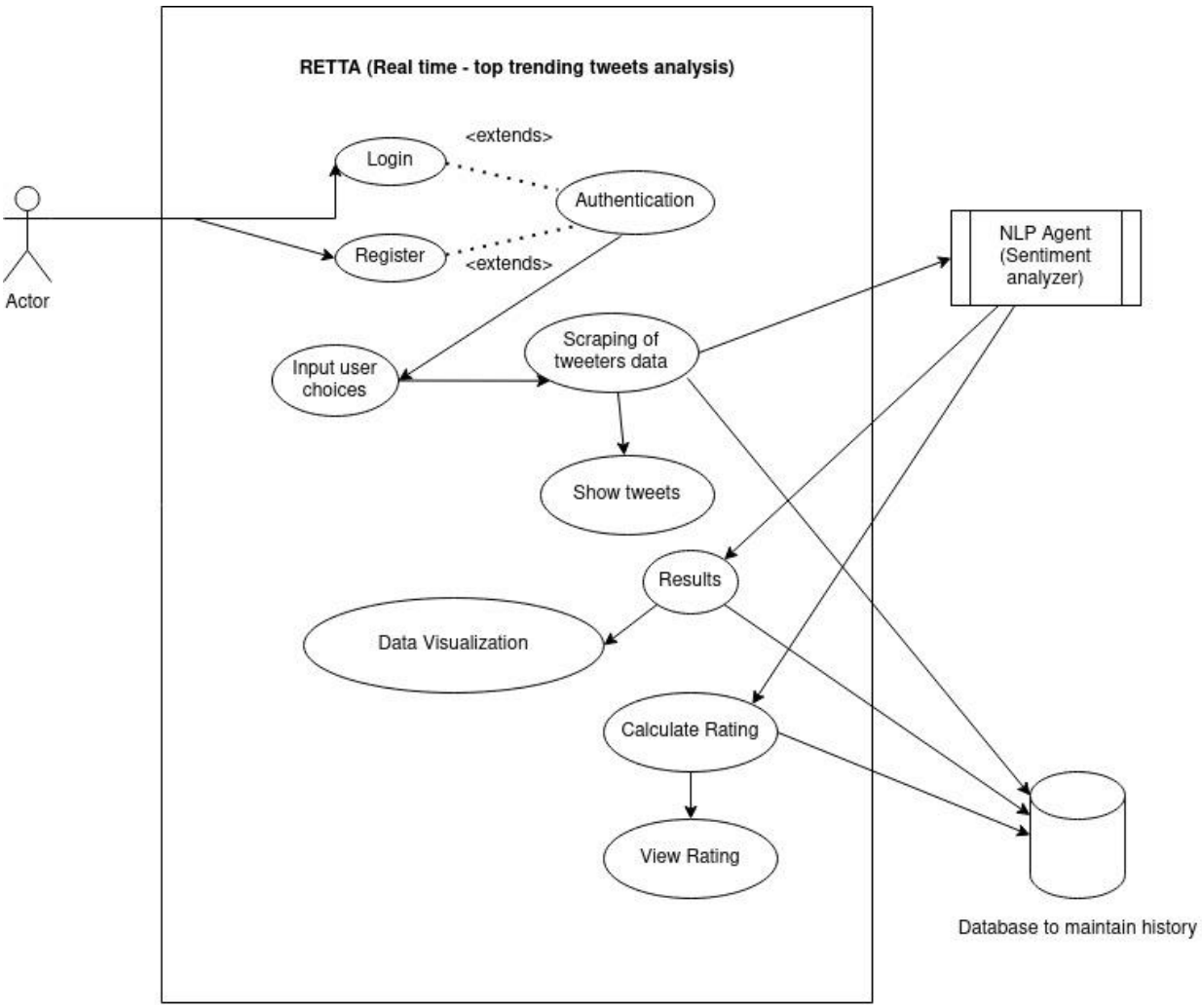


Figure 4 – Use-Case Diagram

10.6 Diagram of Class

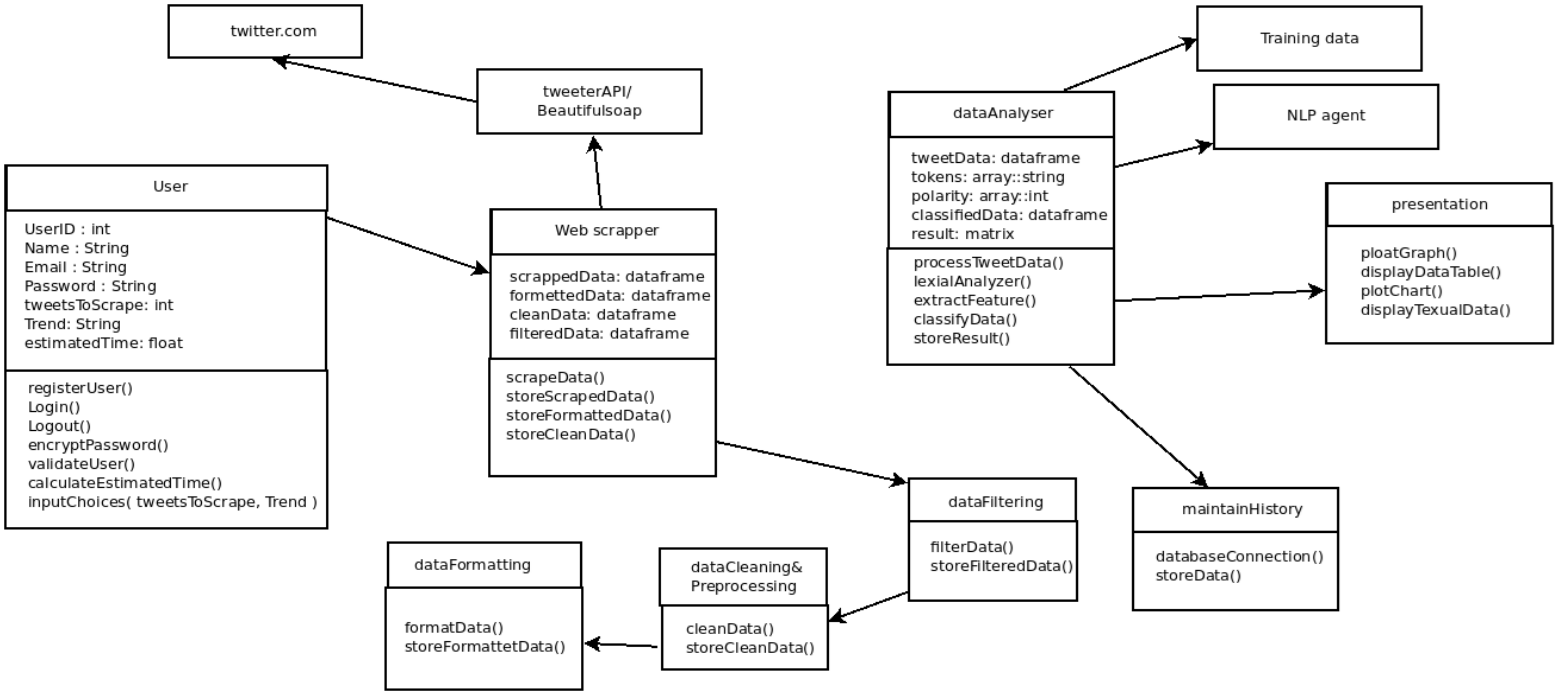


Figure 5 – Class Diagram

Class Name	**Description**
Web Scrapper	This class will scrape real time tweeter's data with the help of tweeter API or some other scrapper like beautiful soap and store formatted and clean data into pandas data-frame.
User	User class contains all the information related to user management. It has aggregation with other classes of user categorization and the functions that performs all the user management functions.
dataFiltering	This is an inherited class of web scrapper, it will filter and store scrapped data into data frame.
DataCleaning	This class is an inherited class of dataFiltering class, it will contain method that can clean formatted data and store it in data frame.
DataAnalyzer	This class contain method like precessTweetData which will process tweet data to extract its features then classify it using NLP agent and training data and then store result

Presentation	This will responsible for data visualization.

10.7 Activity Diagram

10.7.1 User Management

The graphic below shows how users log into the system and are presented with different interfaces depending on their category.

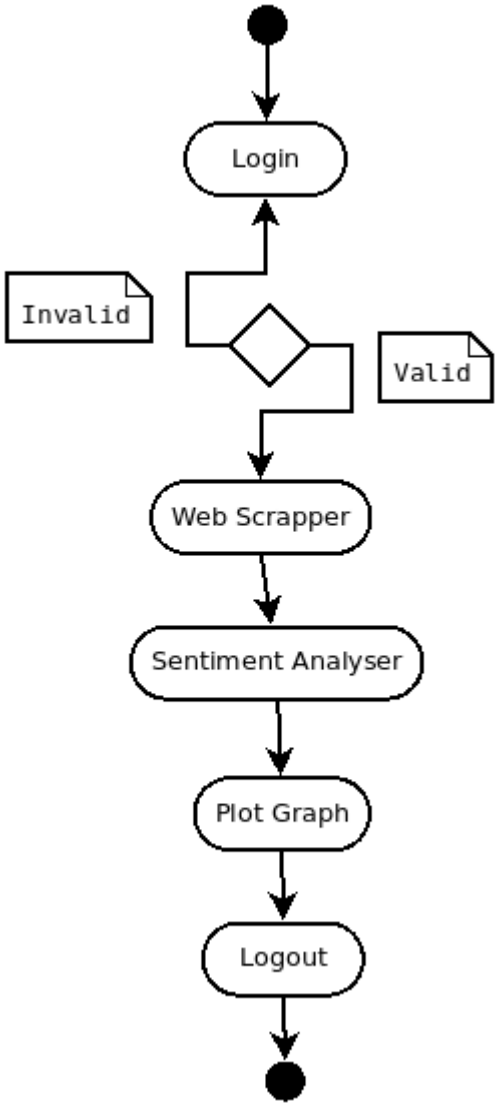


Figure 6 – User Activity Diagram

10.7.1 Web Scrapper

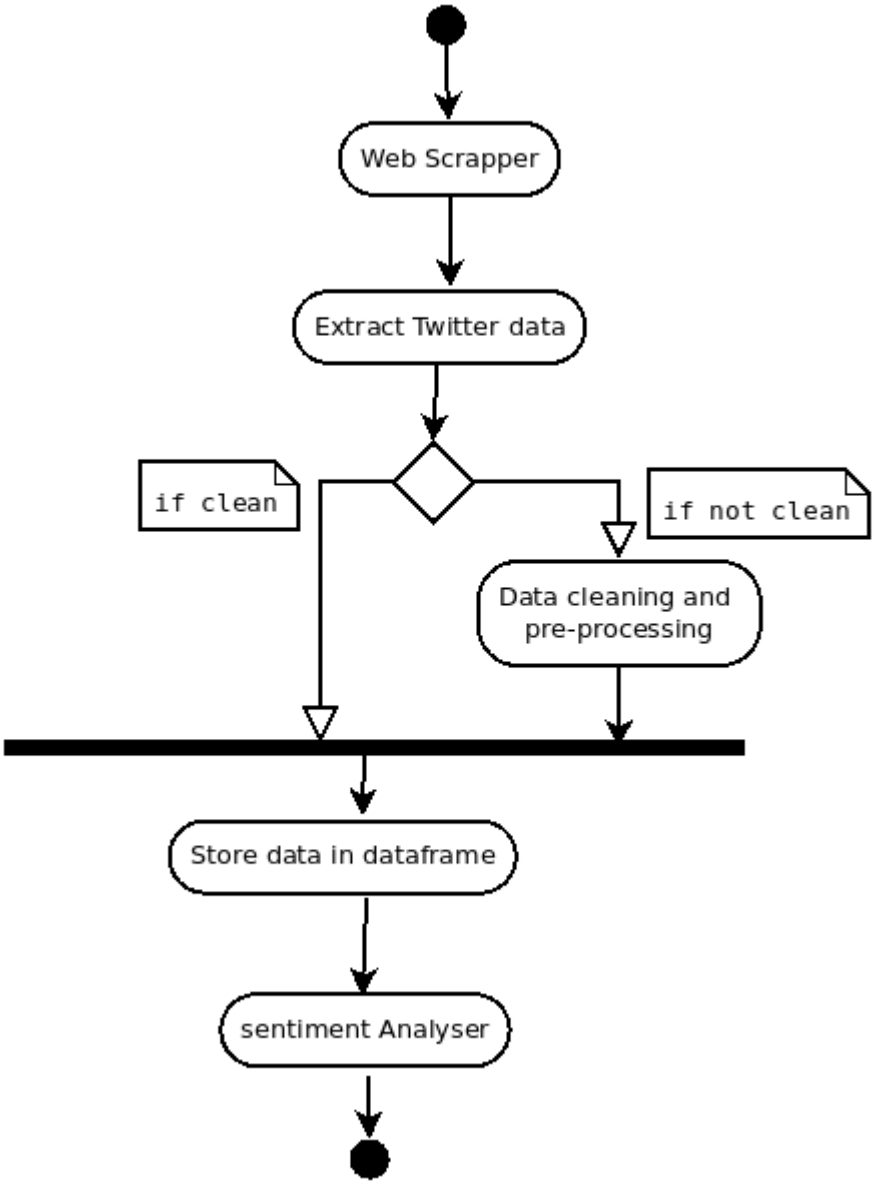


Figure 6.1 – Web Scrapper Activity Diagram

10.7.2 Data Analyzer

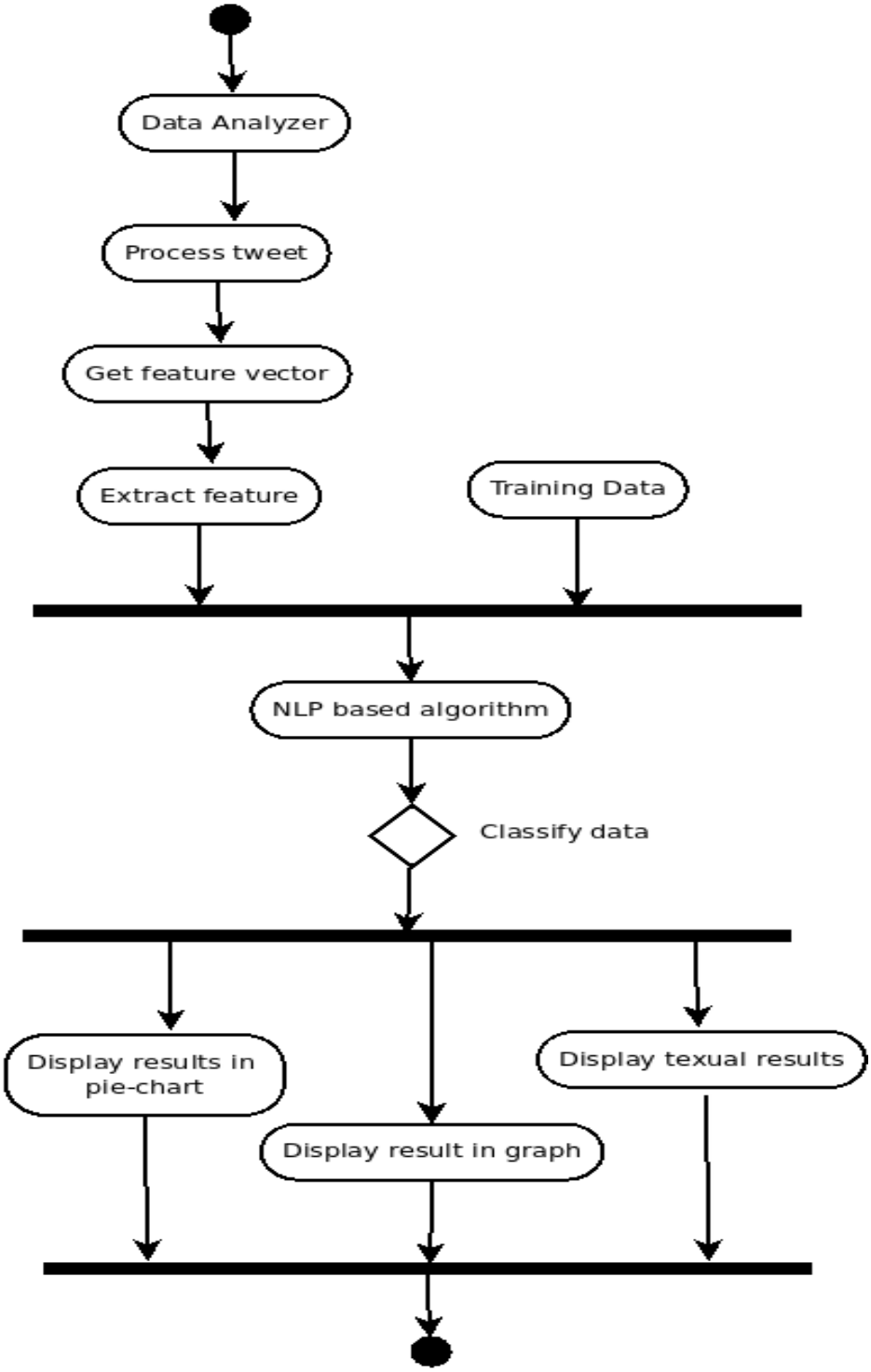


Figure 6.2 – Data Analyzer Activity Diagram

10.8 Sequence Diagram (Dynamic View)

10.8.1 Login

The following diagram shows the sequence of events as a user log in into the system.

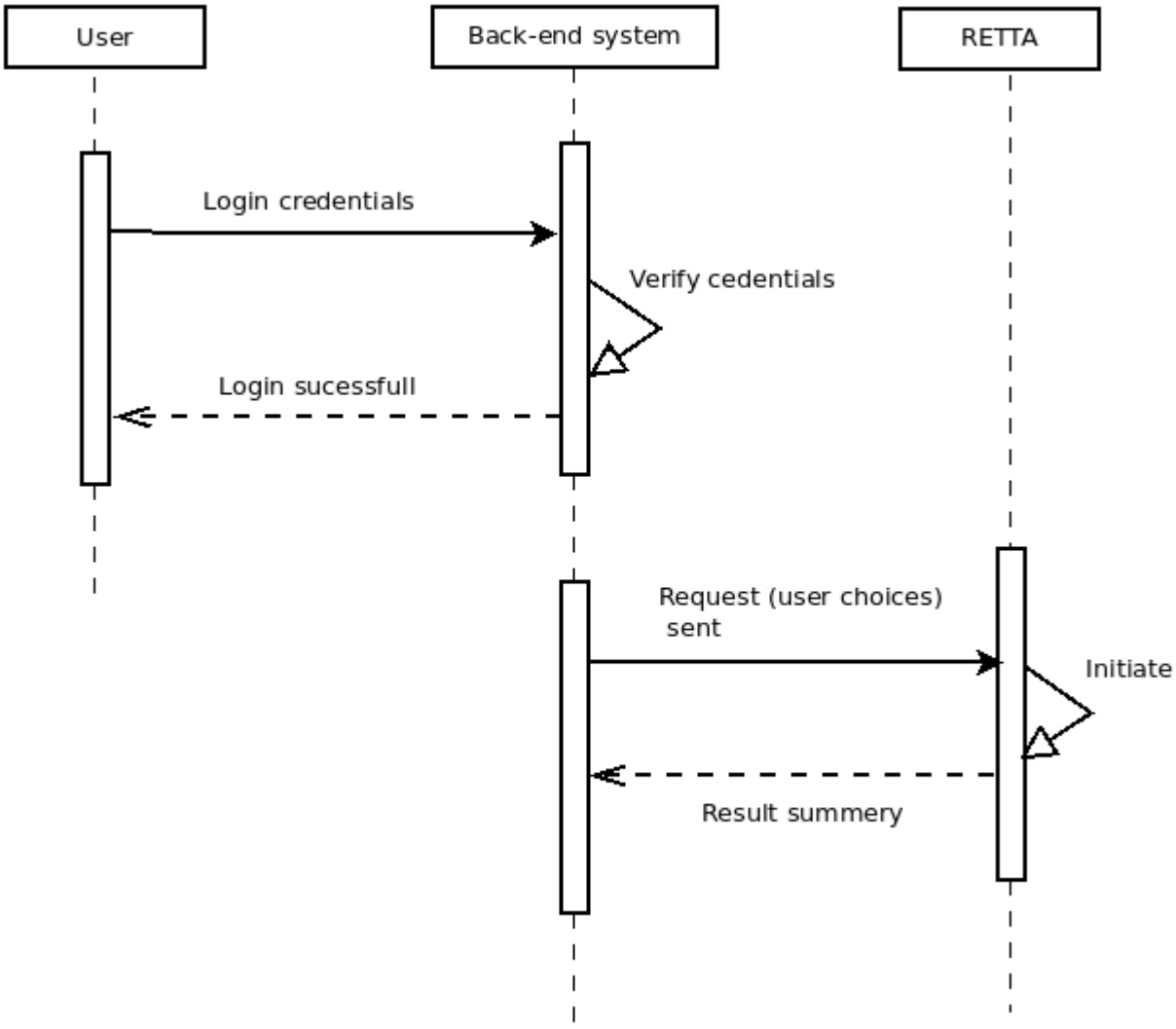


Figure 7 – Login Sequence Diagram

10.8.2 Web Scrapper

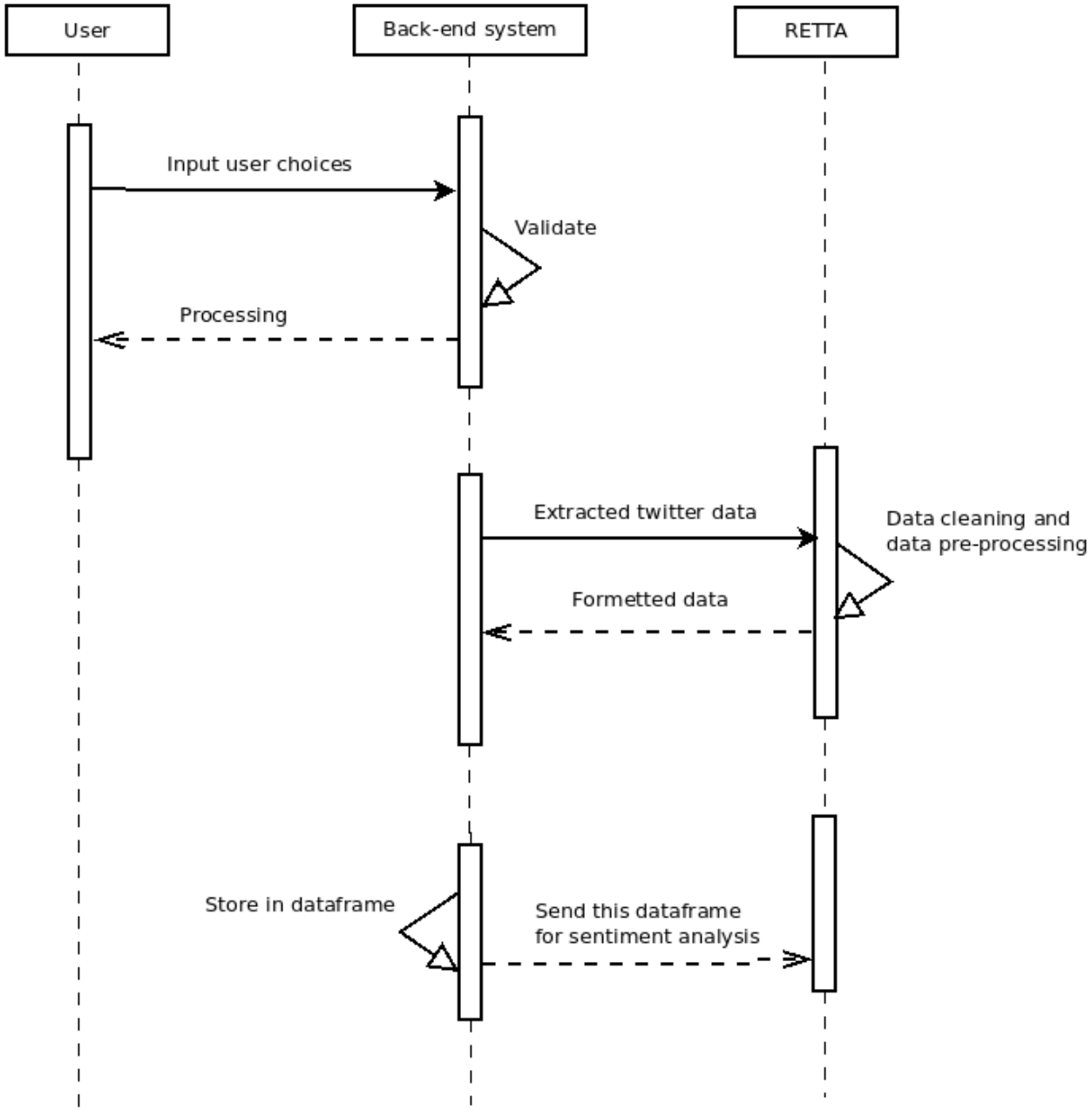


Figure 7.1 – Web Scrapper Sequence Diagram

10.8.3 Data Analyzer

All the effort is focused to analyze data. Following Sequence diagram depicts the procedure to do sentiment analysis of formatted data.

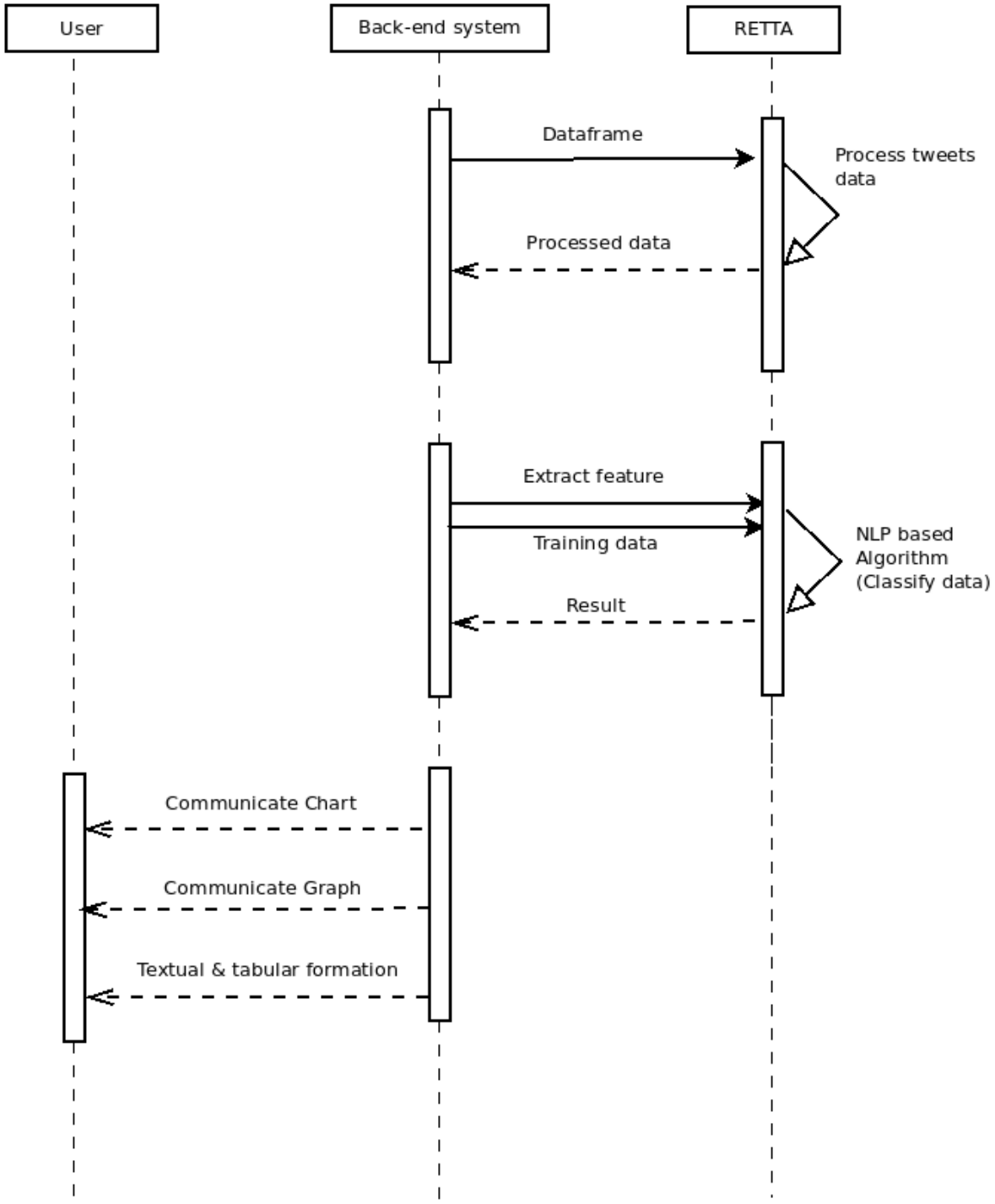


Figure 7.2 – Data Analyzer Sequence Diagram

10.9 User-Interface Design

The system under construction will feature a basic and simple user interface with a very short learning curve and little training to function at optimal efficiency.

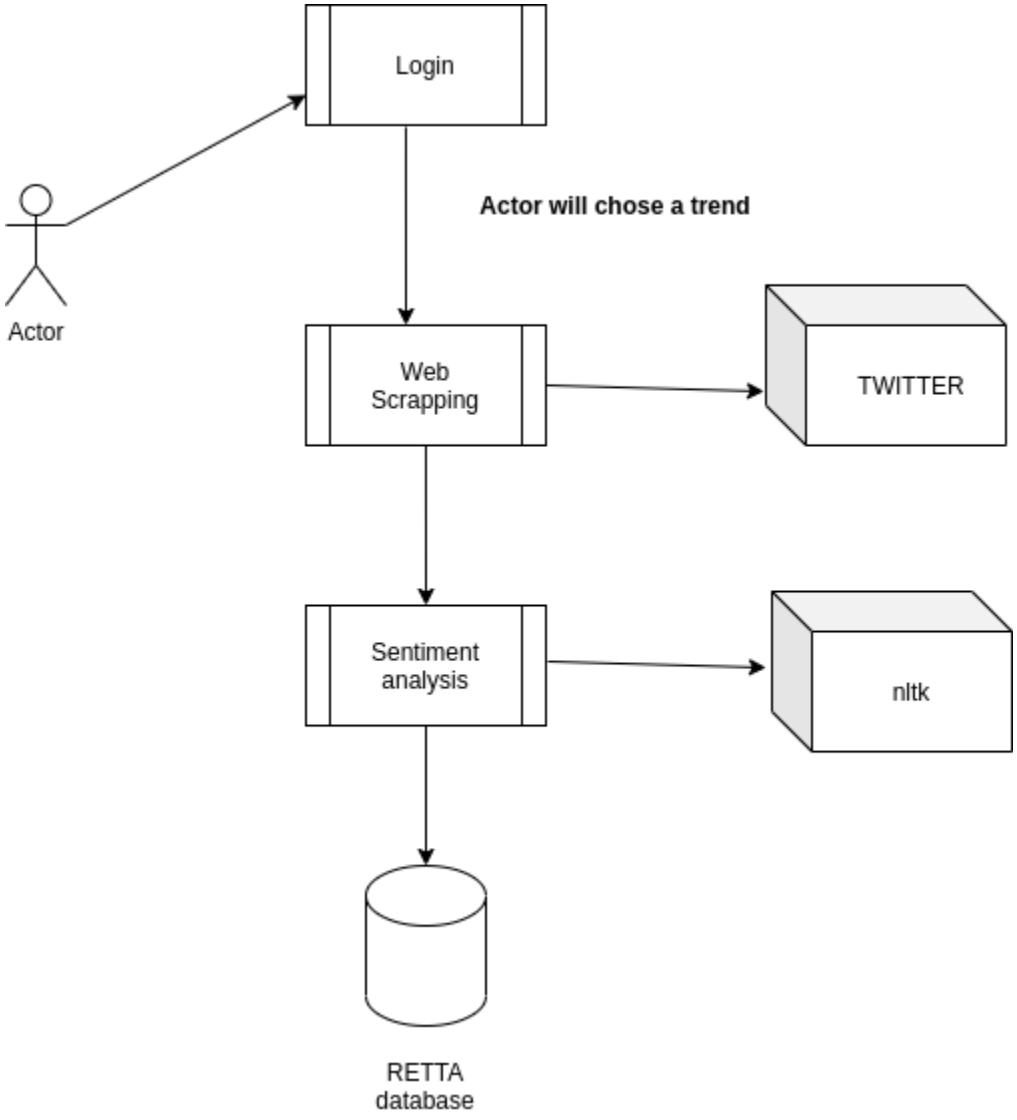
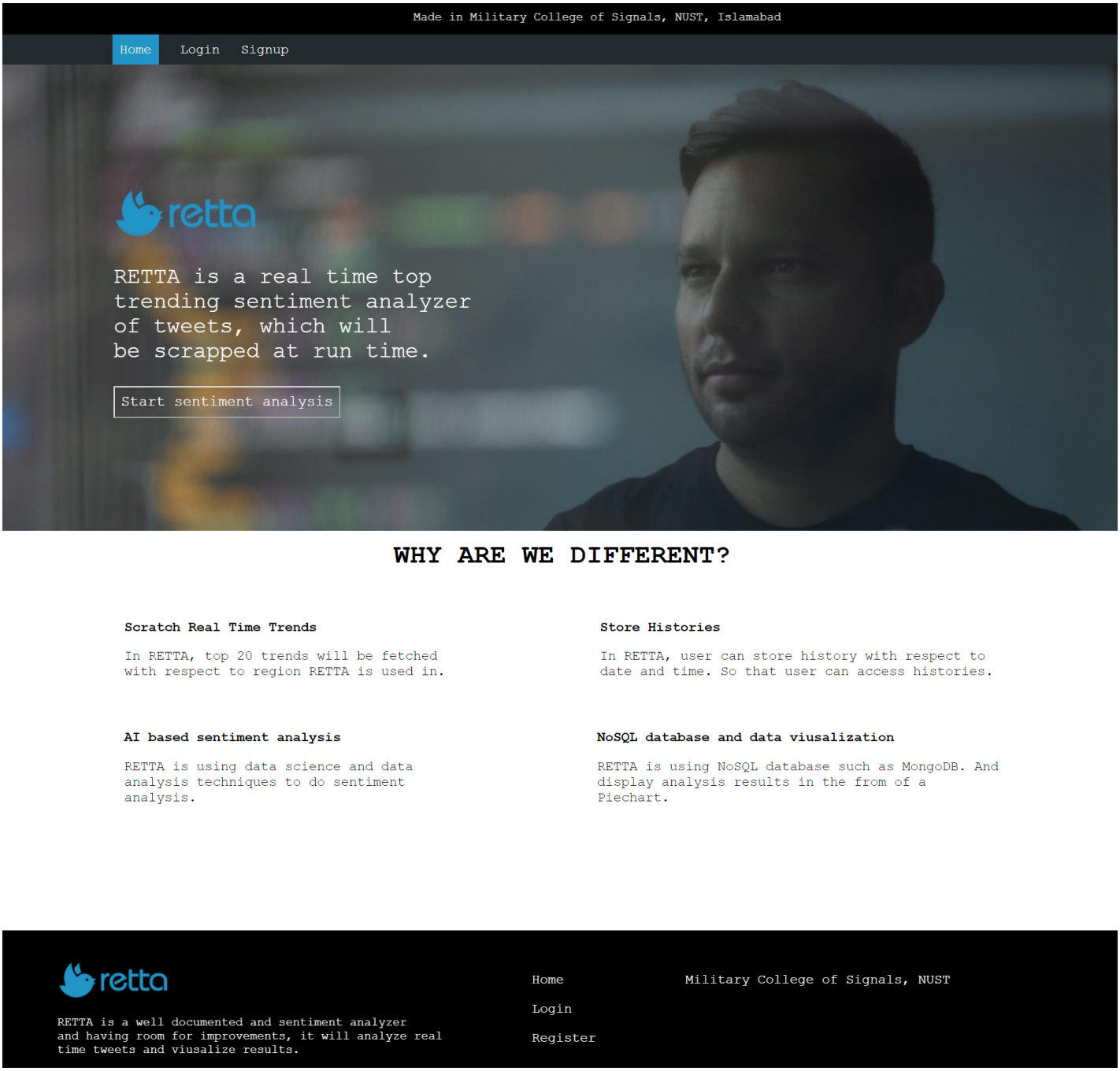


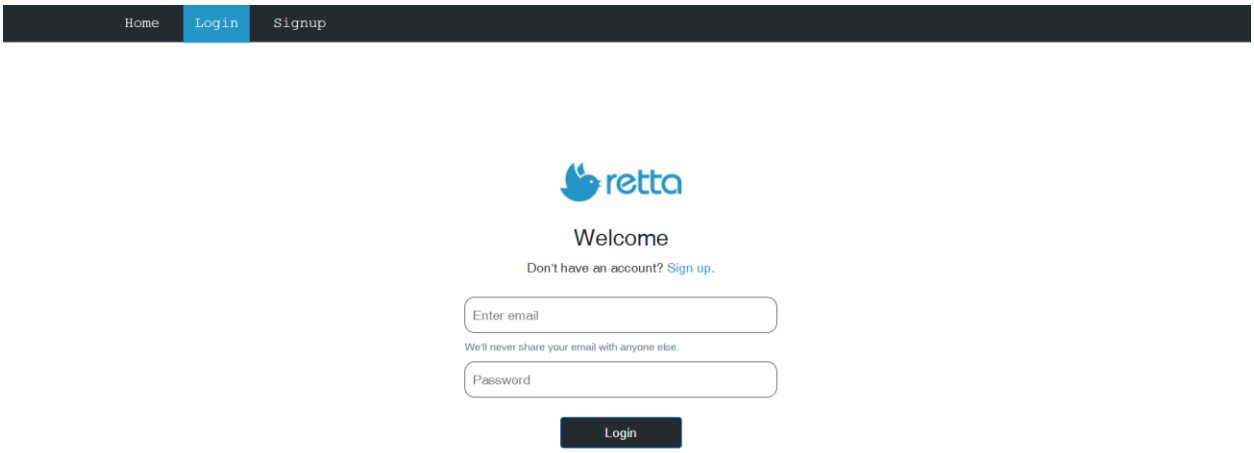
Figure 8 –User-Interface Diagram

10.9.1 Description of Diagram

10.9.1.1 Login

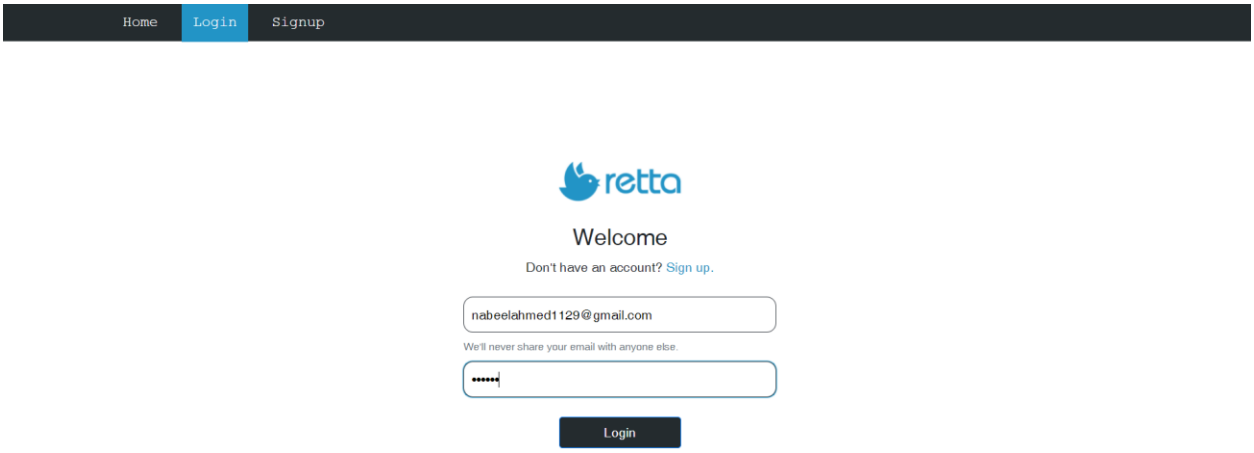
This is the initial login screen, which takes username and password as input. After a successful login, the user is sent to the main interface.

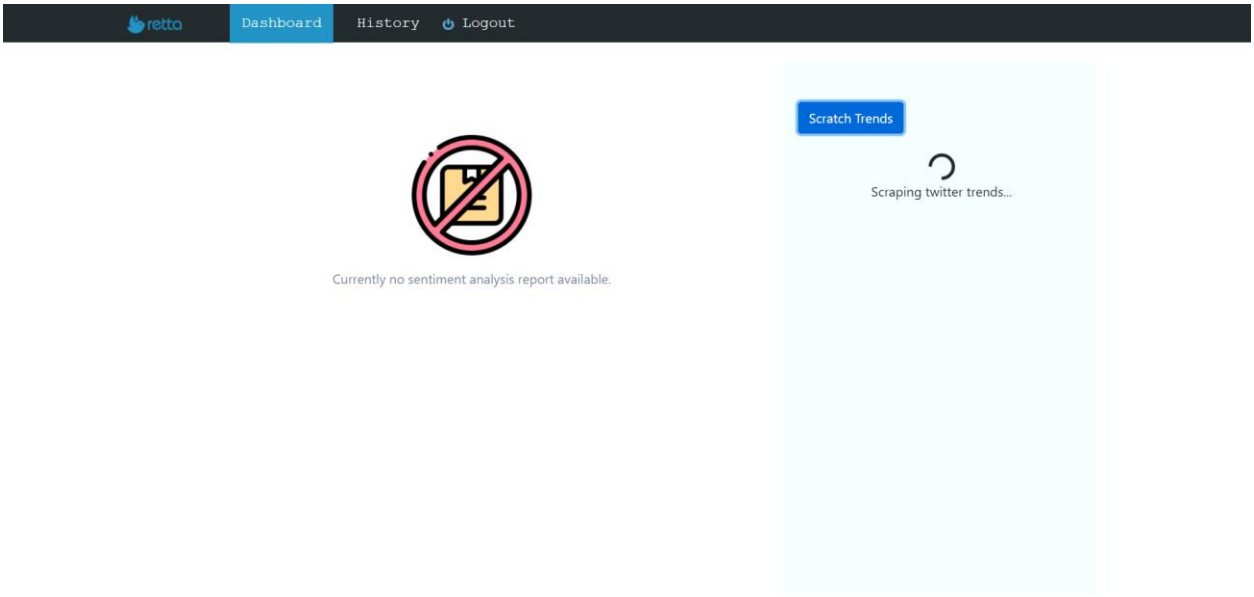


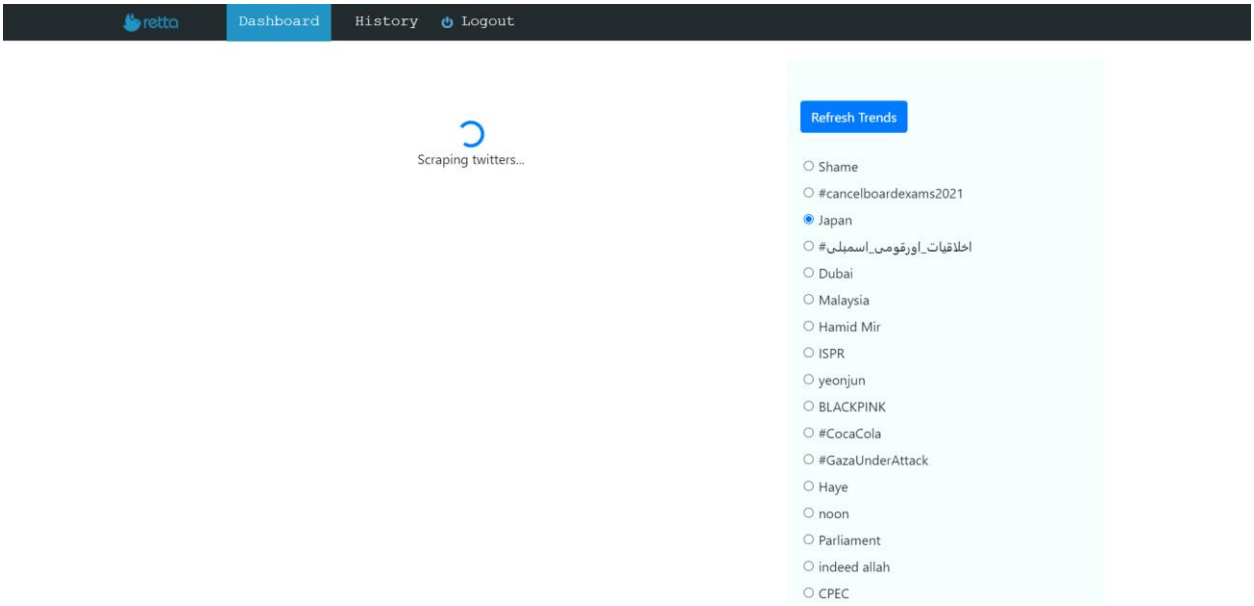


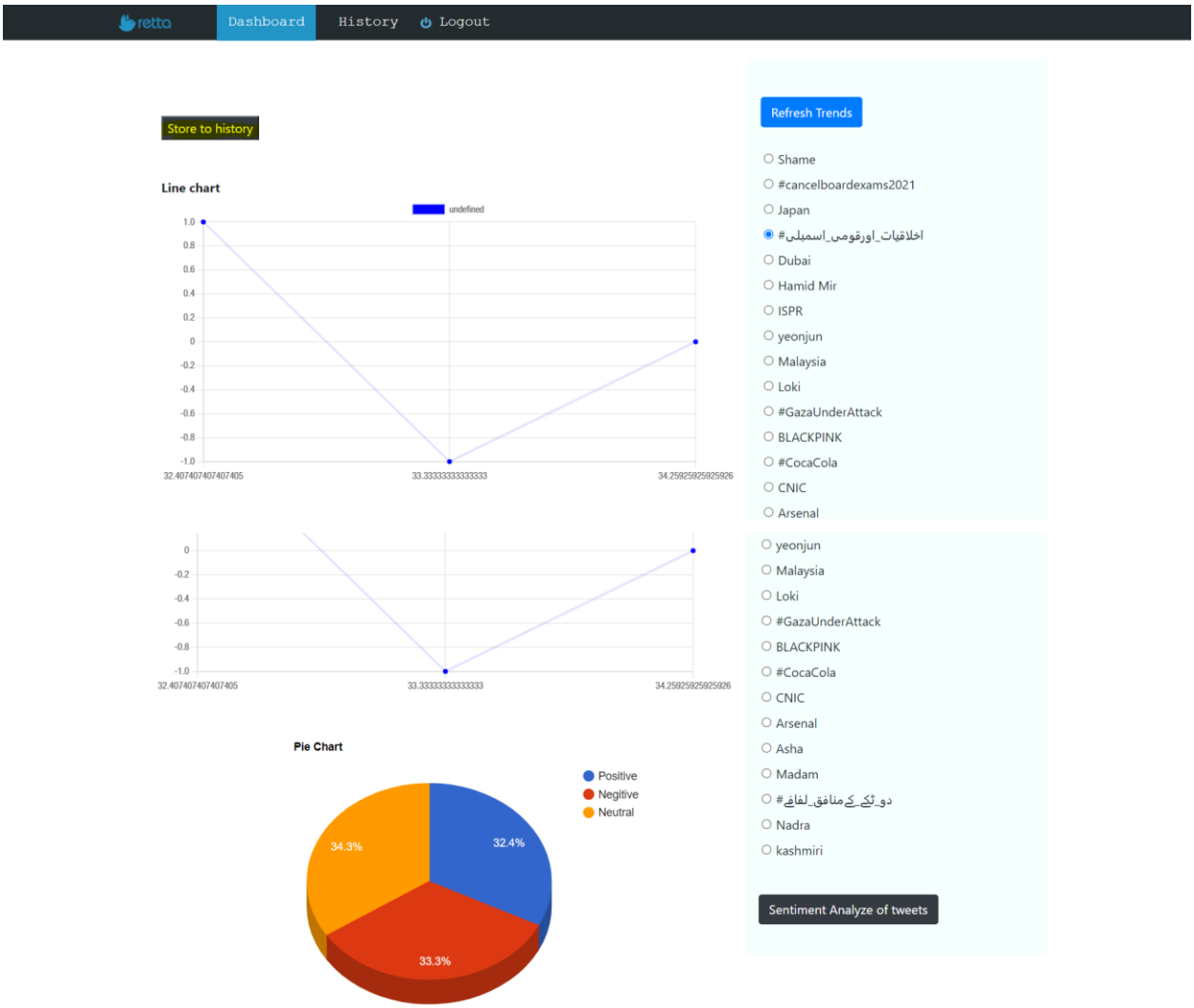
10.9.1.2 Login Screen

Following are the sketches of a possible UI implementation for RETTA. This will be the first screen that the user see upon opening the RETTA. The user shall be presented with a form to enter username and password for login. As visible, the interface is simple and self-explanatory.




retta
Dashboard
History
Logout
Currently no sentiment analysis report available.
Scratch Trends
Scraping twitter trends...


retta
Dashboard
History
Logout
Scraping twitters...
Refresh Trends
Shame
#cancelboardexams2021
Japan
اخلاقیات_اورقومی_اسمبلی#
Dubai
Malaysia
Hamid Mir
ISPR
yeonjun
BLACKPINK
#CocaCola
#GazaUnderAttack
Haye
noon
Parliament
indeed allah
CPEC


retta
Dashboard
History
Logout
Store to history
Line chart
undefined
1.0
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1.0
32.407407407407405
33.33333333333333
34.25925925925926
0
-0.2
-0.4
-0.6
-0.8
-1.0
32.407407407407405
33.33333333333333
34.25925925925926
Pie Chart
Positive
Negitive
Neutral
34.3%
32.4%
33.3%
Refresh Trends
Shame
#cancelboardexams2021
Japan
#اخلاقیات_اورقومی_اسمبلی
Dubai
Hamid Mir
ISPR
yeonjun
Malaysia
Loki
#GazaUnderAttack
BLACKPINK
#CocaCola
CNIC
Arsenal
yeonjun
Malaysia
Loki
#GazaUnderAttack
BLACKPINK
#CocaCola
CNIC
Arsenal
Asha
Madam
#دو_ٹکے_کے_منافق_لفافے
Nadra
kashmiri
Sentiment Analyze of tweets

retta Dashboard History Logout

Histories

Date	Trend	Positive (%)	Negitive (%)	Neutral (%)
Wed Jun 09 2021 21:09:30 GMT+0500 (Pakistan Standard Time)	#HEATWAVE	29.694	13.537	56.769
Wed Jun 09 2021 21:11:36 GMT+0500 (Pakistan Standard Time)	#PSL6	25.749	2.395	71.856
Wed Jun 09 2021 21:15:18 GMT+0500 (Pakistan Standard Time)	MBBS	38.060	23.881	38.060
Wed Jun 09 2021 22:06:08 GMT+0500 (Pakistan Standard Time)	#HEATWAVE	33.772	12.719	53.509
Wed Jun 16 2021 14:13:04 GMT+0500 (Pakistan Standard Time)	Japan	63.462	3.846	32.692

Before	For Human	For ML model
Sunni Shahzada@SunniShahzada2·15hDo not test someone by his dress but by his manners and speech. #اخلاقیات_اورقومی_اسمبلی12152290	Do not test someone by his dress but by his manners and speech#اخلاقیات_اورقومی_اسمبلی.	test someon dress manner speech. اخلاقیاتاورقومیاسمبلی
Rashid Khan@sa_raki·15hSaad Hussain Rizvi is the name of a love. And love never dies #2:1914اخلاقیات_اورقومی_اسمبلیK viewsFrom Rashid Khan7212319	Saad Hussain Rizvi is the name of a love. And love never dies #2:1914اخلاقیات_اورقومی_اسمبلیK viewsFrom Rashid Khan	saad hussain rizvi name love love never die 21914اخلاقیاتاورقومیاسمبلیk viewsfrom rashid khan
Tahreak Labiak Pakistan PP39 (Official)@OfficialTLPPP39·15hBe Attentive #اخلاقیات_اورقومی_اسمبلی Trend For Today23261361	Be Attentive#اخلاقیات_اورقومی_اسمبلی Trend For Today	attentiveاخلاقیاتاورقومیاسمبلی trend today
KaShMiRi RaJpOot@KHR_Aati·15hThe moral is that when it comes to your caste, forgive and when it comes to the honor of the Messenger of Allah, cut it off. #2119226اخلاقیات_اورقومی_اسمبلی	The moral is that when it comes to your caste, forgive and when it comes to the honor of the Messenger of Allah, cut it off.#اخلاقیات_اورقومی_اسمبلی	moral come cast forgiv come honor messeng allah cut offاخلاقیاتاورقومیاسمبلی
Ayaansudeas@Hashir_shanji·15hFauzia Kasuri, a founding member of PTI and former Tai mother of youth, who came	Fauzia Kasuri, a founding member of PTI and former Tai mother of youth, who came to her senses quickly. I pray for the rest that	fauzia kasuri found member pti former tai mother youth came sens quickli pray rest ثیاپلیپیروموشناخلاقیاتاورقومیاسمبلی

Top 100 tweets sample report

Tweet	Sentiment	Polarity
,,,Everyone says now No more no more,,,,#اخلاقیات_اورقومی_اسمبلی	NEUTRAL	0
Say the wrong thing wrong Because history is written about collisions Not of those who lick the soles#TeamRRF#اخلاقیات_اورقومی_اسمبلی	NEGATIVE	-0.7351
A commotion in the House, can the Speaker take legal action? 'Democracy has only one dictator, and he is the speaker' Surprised at his silence.#اخلاقیات_اورقومی_اسمبلی	POSITIVE	0.128
The Modern Monk@bookworm_monk·54mCristiano Ronaldo snub wipes billions off Coca-Cola's market value#CocaCola #CristianoRonaldo #0:443اخلاقیات_اورقومی_اسمبلیK views	NEGATIVE	-0.4215
Saad Hussain Rizvi is the name of a love. And love never dies #2:1914اخلاقیات_اورقومی_اسمبلیK viewsFrom Rashid Khan	POSITIVE	0.9108
Revolution has become the destiny of this country. Today, the heads of the entire nation have bowed in shame over the filthy insults uttered in the parliament. @Ati_gujr#اخلاقیات_اورقومی_اسمبلی#ٹی_ایل_پی_پروموش	NEGATIVE	-0.7506
In the National Assembly today, all ethics were completely violated#اخلاقیات_اورقومی_اسمبلی	NEUTRAL	0
The educated class showed its European training today#اخلاقیات_اورقومی_اسمبلی	NEUTRAL	0

CHAPTER 4

QUALITY ASSURANCE

CHAPTER 4: QUALITY ASSURANCE

The relevant strategies, processes, and techniques utilized to design, execute, and manage testing of the "Real Time top trending tweet analysis" are described in this test plan document. The test strategy will verify that the application's criteria and standards are met to an approved level by the client.

Manual testing will be used, which entails testing software without the use of any automated tools or scripts. In this scenario, the tester assumes the role of an end-user and tests the application for any abnormal behavior or bugs. All functional, application performance, and use case criteria mentioned in the requirement document are covered by the test scope. Software testing can be done at any point during the development phase, depending on the testing approach used. However, after the requirements have been developed and the coding process has been done, the majority of the testing work happens.

11 Test Items

- Development of test scenarios
- Execute multiple tests based on the above-mentioned test scenarios that have been generated
- Inform the appropriate developer or management about any bugs
- Develop and provide test results
- Manage or incorporate adjustments at a later stage in the project's development.

11.1 Features Required to be Tested

Following features are tested:

- Software will be able to login from any selected stand-alone PC after logging secret password.
- After that user will click a button to scrape top twenty trends.
- Then select a trend for which wants a sentiment analysis, click the next button
- After sometime a pie chart will appear on the screen that show percentage of positive, negative and neutral.

- One report is shown in tabular form that show pre-processing result.

Another report is shown that illustrate top 100 tweets and their sentiments against each tweet.

11.2 Test Approaches

This acceptance test plan will be followed throughout the execution of the acceptance test. After all test cases have been completed, a test report will be generated to demonstrate the simulator's quality in providing a meaningful learning application. In order to execute the test, the following test techniques will be used:

Unit Testing: Unit testing is the responsibility of developers. Each module's and component's implementation will be checked individually.

Integration Testing: The integration test cases will be executed when the unit test has passed over the chosen quality level. It's critical to test the product as a black-box after all of the modules have been integrated.

Positive and Negative Testing: This method will be used in conjunction with unit and integration testing. Test cases are written in scenarios that are evident and guarantee that all functional criteria are met. Furthermore, many test cases will be presented to demonstrate how the program responds to invalid operations.

11.3 Pass or Fail Measure

Details of the test cases are specified in section Test Deliverables. Following the principles outlined below, a test item would be judged as pass or fail.

- Pre-conditions fulfilled
- Inputs carried out according to plan
- The output matches what was indicated in output => Pass
- The system does not function or does not meet the output requirements => Fail

11.4 Standard for Deferral and Renewal Requirements

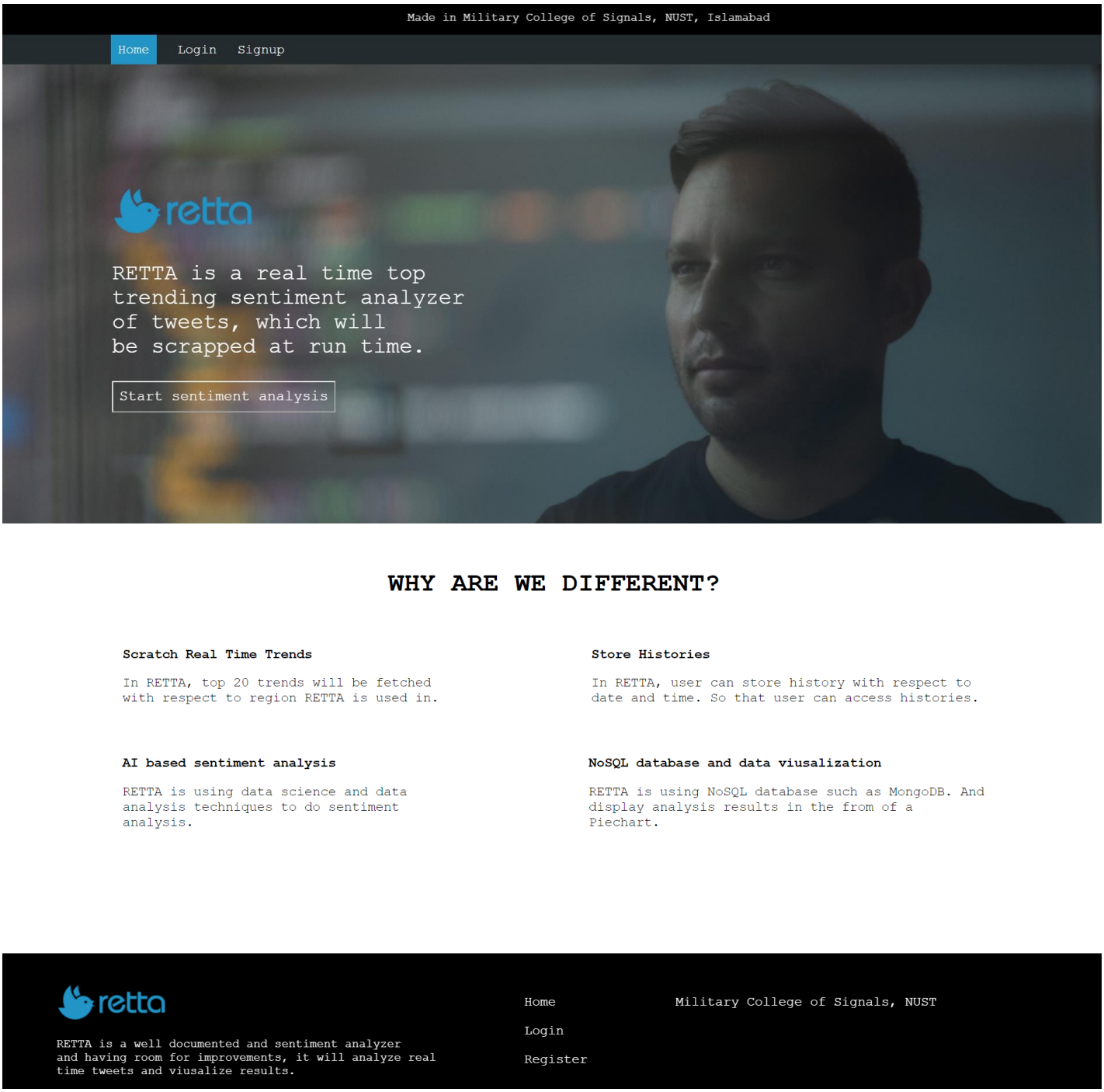
Developers can rapidly correct any flaws discovered, eliminating the need to restart the testing process from the beginning. However, when serious flaws prevent certain test cases from running because they are interdependent, testing must be suspended.

11.5 Test Deliverables

Following are the Test Cases.

11.5.1 User Interface Testing

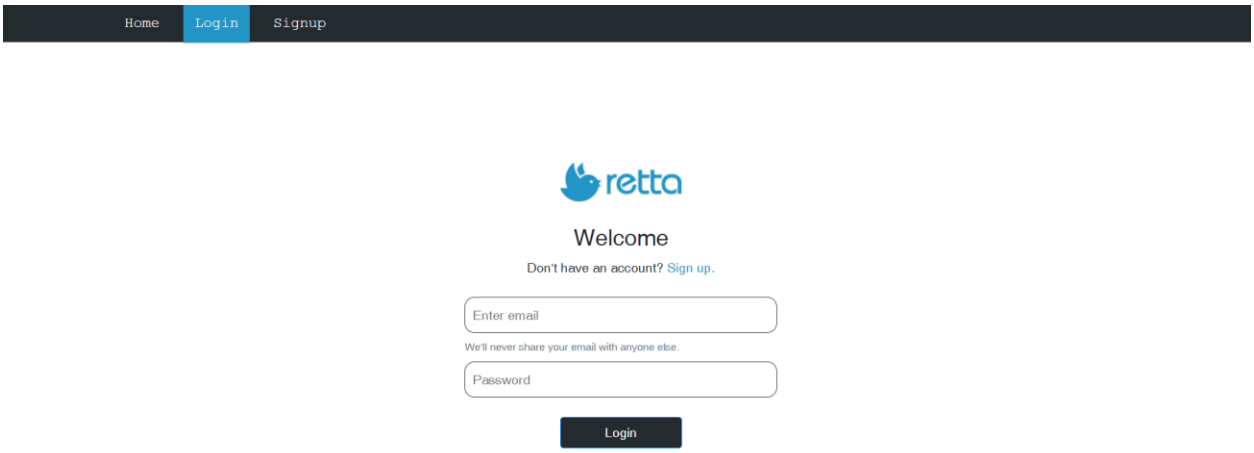
Test-Case No	1
Test Case Name	Open RETTA
Description	Testing Application whether it runs on web browser or not
Testing Technique	Unit testing, Black Box Testing
Preconditions	Chrome browser will be also on your system and al chrome driver
Input Values	User name and password
Steps	• Click on RETTA icon. • Enter username and password.



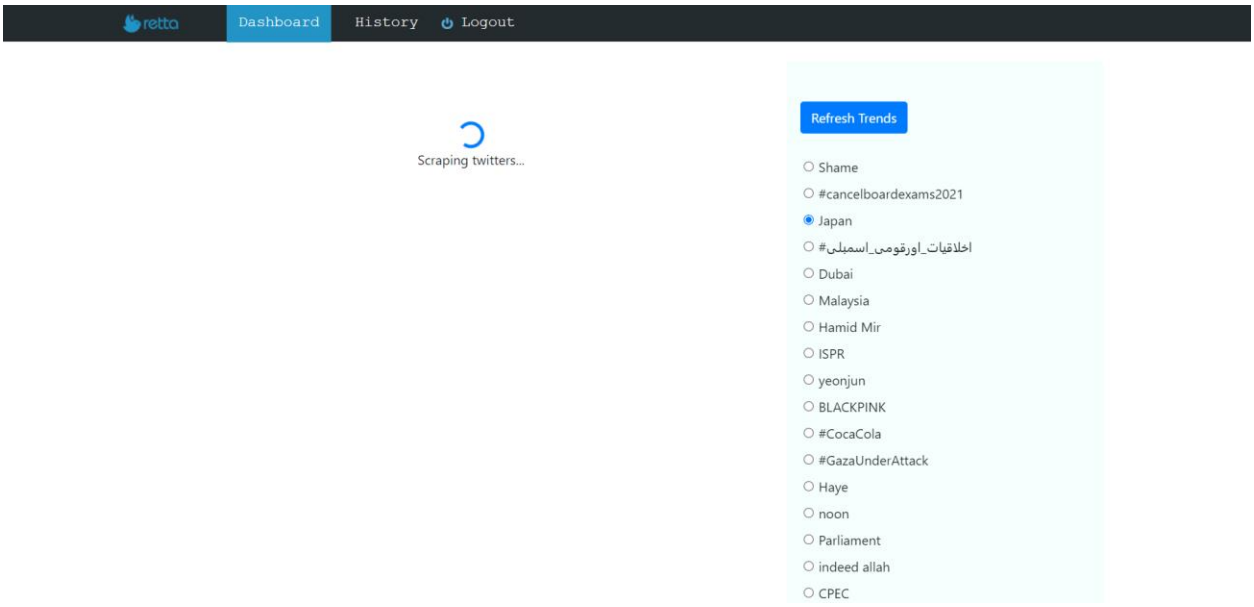
Test Case Number	02
Test Case Name	Open RETTA
Description	Testing Application whether it runs on PC or not.
Testing Technique	Unit testing, Black Box Testing

Preconditions	Chrome browser will be also on your system and al chrome driver
Input Values	Incorrect User name/ password
Steps	• Click on RETTA icon. • Enter username and password. • Click login button.
Expected output	RETTA main menu will not open.
Actual output	Error message generated.
Status	Test case passed successfully.

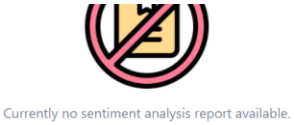
Test-Case No.	03
Test-Case Name	Database Case
Description	Testing function by Maintaining/Registering database of users.
Testing Technique	Unit testing, Black Box Testing
Preconditions	Chrome browser will be also on your system and al chrome driver
Input Values	Service number, Name ,Unit, Formation, Age
Steps	• Click on RETTA icon. • Click on register. • Enter particulars. • Save particulars
Expected output	Interface should be displayed with successful message .
Actual output	Message displayed successfully
Status	Test case passed successfully.

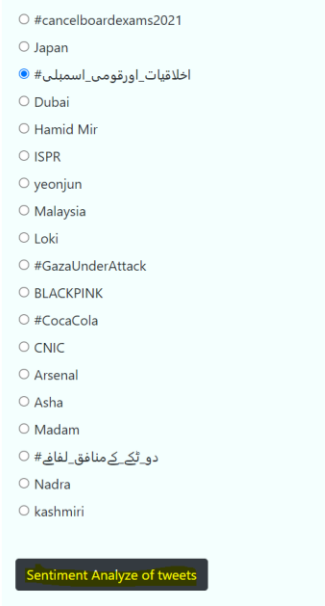


Test Case Number	04
Test Case Name	Twitter Scrapping
Description	Press button to scrape tweets from twitter
Testing Technique	Unit testing, Black Box Testing
Preconditions	Application is running with authentic login
Input Values	Click on " Scrape tweets " Button
Steps	• Run the application • Login page is open. • Enter password. • Click on “Scrape tweets” button.
Expected output	Top twenty tweets will be shown on screen.
Actual output	Top twenty tweets are shown on twitter.
Status	Test case passed successfully.

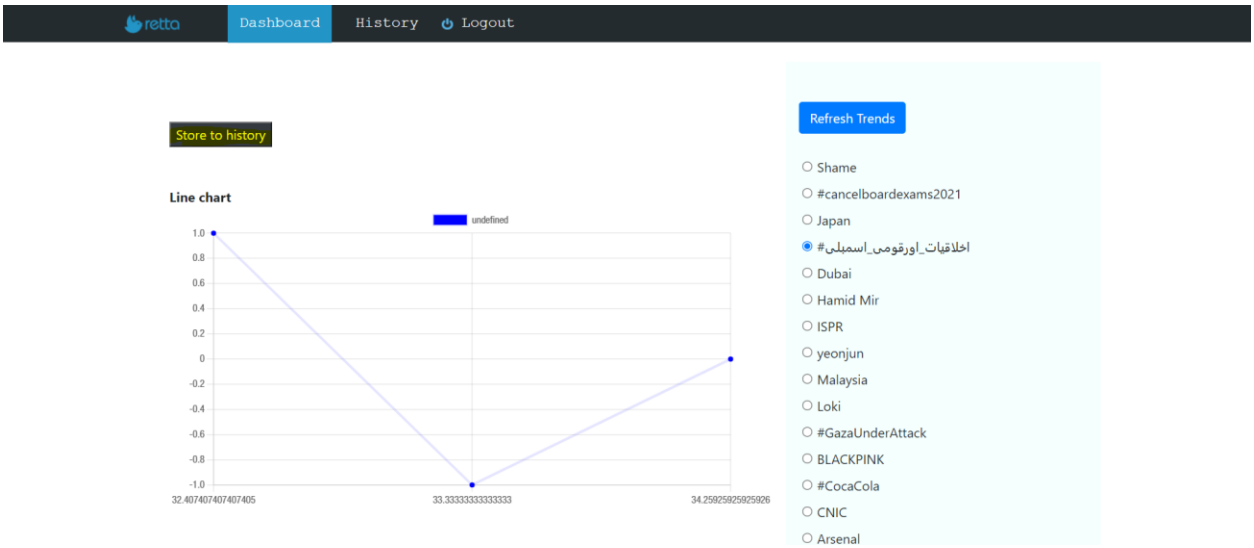


Test Case Number	05
Test Case Name	Tweet selection
Description	Select a tweets from a list of tweets
Testing Technique	Unit testing, Black Box Testing
Preconditions	Application is running with authentic login
Input Values	Click on " Scrape tweet" Button
Steps	• Run the application • Login page is open. Enter password. • Select a single tweet from a list • Click on “Scrape tweet” button.
Expected output	Pie chart will be displayed
Actual output	Pie chart will be displayed.
Status	Test case passed successfully.

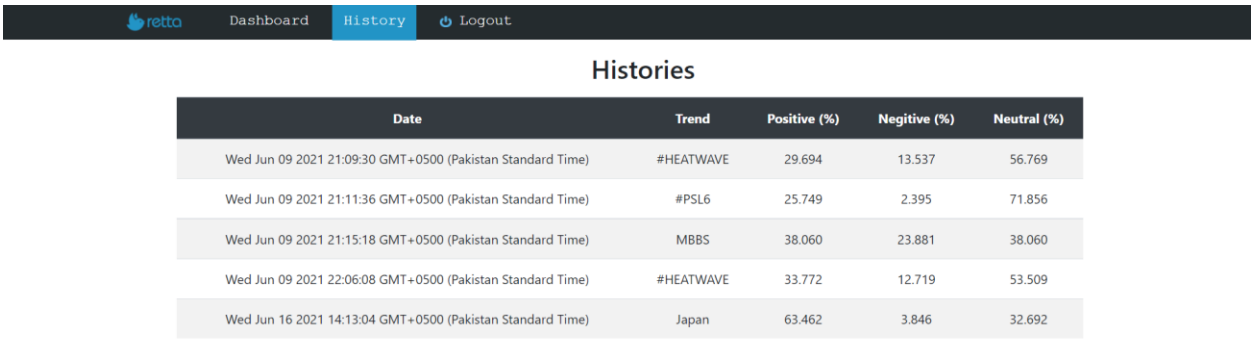




Test Case Number	06
Test Case Name	Store History
Description	Press button to store history
Testing Technique	Unit testing, Black Box Testing
Preconditions	Application is running with authentic login
Input Values	Click on " Store history " Button
Steps	• Run the application • Login page is open. Enter password. • Select a trend and get a pie chart • Click Store history button.
Expected output	Success message will be shown
Actual output	Alert shown with success messages
Status	Test case passed successfully.



Test Case Number	07
Test Case Name	Show Histories
Description	Press show History button to show Histories that user have stored
Testing Technique	Unit testing, Black Box Testing
Preconditions	Application is running with authentic login
Input Values	Click on " Histories " Button to show all histories
Steps	• Run the application • Login page is open. • Enter password. • Click on Histories button • Screen with the histories shown.
Expected output	Screen with the histories shown
Actual output	Page displayed
Status	Test case passed successfully.



Test Case Number	08
Test Case Name	Prepossessing report will be shown
Description	Preprocessing report will be shown in the form of table
Testing Technique	Unit testing, Black Box Testing
Preconditions	Application is running with authentic login
Input Values	It will come along pie chart
Steps	• Run the application • Login page is open. • Enter password. • Select tweet and get scrapped result.
Expected output	Report will be shown on the screen n the form of table
Actual output	Report will be shown on the screen n the form of table
Status	Test case passed successfully.

Before	For Human	For ML model
Sunni Shahzada@SunniShahzada2·15hDo not test someone by his dress but by his manners and speech. 12152290#اخلاقیات_اورقومی_اسمبلی	Do not test someone by his dress but by his manners and speech#.اخلاقیات_اورقومی_اسمبلی	test someon dress manner speech. اخلاقیاتاورقومیاسمبلی
Rashid Khan@sa_raki·15hSaad Hussain Rizvi is the name of a love. And love never dies #2:1914اخلاقیات_اورقومی_اسمبلیK viewsFrom Rashid Khan7212319	Saad Hussain Rizvi is the name of a love. And love never dies #2:1914اخلاقیات_اورقومی_اسمبلیK viewsFrom Rashid Khan	saad hussain rizvi name love love never die 21914اخلاقیاتاورقومیاسمبلیk viewsfrom rashid khan
Tahreak Labiak Pakistan PP39 (Official)@OfficialTLPPP39·15hBe Attentive #اخلاقیات_اورقومی_اسمبلی Trend For Today23261361	Be Attentive#اخلاقیات_اورقومی_اسمبلی Trend For Today	attentiveاخلاقیاتاورقومیاسمبلی trend today
KaShMiRi RaJpOot@KHR_Aati·15hThe moral is that when it comes to your caste, forgive and when it comes to the honor of the Messenger of Allah, cut it off. #2119226اخلاقیات_اورقومی_اسمبلی	The moral is that when it comes to your caste, forgive and when it comes to the honor of the Messenger of Allah, cut it off.#اخلاقیات_اورقومی_اسمبلی	moral come cast forgiv come honor messeng allah cut offاخلاقیاتاورقومیاسمبلی
Ayaansudeas@Hashir_shanji·15hFauzia Kasuri, a founding member of PTI and former Tai mother of youth, who came	Fauzia Kasuri, a founding member of PTI and former Tai mother of youth, who came to her senses quickly. I pray for the rest that	fauzia kasuri found member pti former tai mother youth came sens quickli pray rest ثیاپلیپیروموشناخلاقیاتاورقومیاسمبلی

Test Case Number	9
Test Case Name	Top 100 tweet report
Description	Top 100 tweets will be shown, with sentiment present along with each tweet
Testing Technique	Unit testing, Random testing Black Box Testing
Preconditions	Application is running with authentic login
Input Values	Come along with pie chart
Steps	• Run the application • Login page is open. • Enter password. • Select trend and scrape tweets. • Shown on the screen
Expected output	Top 100 tweets will be shown on the screen.
Actual output	Top 100 tweets will be shown on the screen.
Status	Test case passed successfully.

Top 100 tweets sample report

Tweet	Sentiment	Polarity
,,,Everyone says now No more no more,,,,#اخلاقیات_اورقومی_اسمبلی	NEUTRAL	0
Say the wrong thing wrong Because history is written about collisions Not of those who lick the soles#TeamRRF#اخلاقیات_اورقومی_اسمبلی	NEGATIVE	-0.7351
A commotion in the House, can the Speaker take legal action? 'Democracy has only one dictator, and he is the speaker' Surprised at his silence.#اخلاقیات_اورقومی_اسمبلی	POSITIVE	0.128
The Modern Monk@bookworm_monk·54mCristiano Ronaldo snub wipes billions off Coca-Cola's market value#CocaCola #CristianoRonaldo #0:443اخلاقیات_اورقومی_اسمبلیK views	NEGATIVE	-0.4215
Saad Hussain Rizvi is the name of a love. And love never dies #2:1914اخلاقیات_اورقومی_اسمبلیK viewsFrom Rashid Khan	POSITIVE	0.9108
Revolution has become the destiny of this country. Today, the heads of the entire nation have bowed in shame over the filthy insults uttered in the parliament. @Ati_gujr#اخلاقیات_اورقومی_اسمبلی#ٹی_ایل_پی_پروموش	NEGATIVE	-0.7506
In the National Assembly today, all ethics were completely violated#اخلاقیات_اورقومی_اسمبلی	NEUTRAL	0
The educated class showed its European training today#اخلاقیات_اورقومی_اسمبلی	NEUTRAL	0

11.6 Risk & Eventualities

11.6.1 Risk Scheduling

The project is on schedule there is no schedule risk.

11.6.2 Risk in Operation

Daily meetings and regular deadlines will be scheduled to accomplish the project's goals and ensure good communication within the group, reducing operational risks.

11.6.3 Risk in Technical Domain

By maintaining the previously specified standards, technical risks will be removed.

11.6.4 Risk associated with Program

In the event of a programming risk, the project's scope will be curtailed in order to keep within the project's restrictions.

CHAPTER 5

USER MANUAL

CHAPTER 5: USER MANUAL

This user guide will help you to get the complete over view how to use this project and details about all the tools that is used. Main tools that are used to develop this project are python 3, twitter, pandas, Numpy Bootstrap, NLTK and selenium.

12 Installation and Usage Guide

Before running RETTA we have to make sure to install following packages

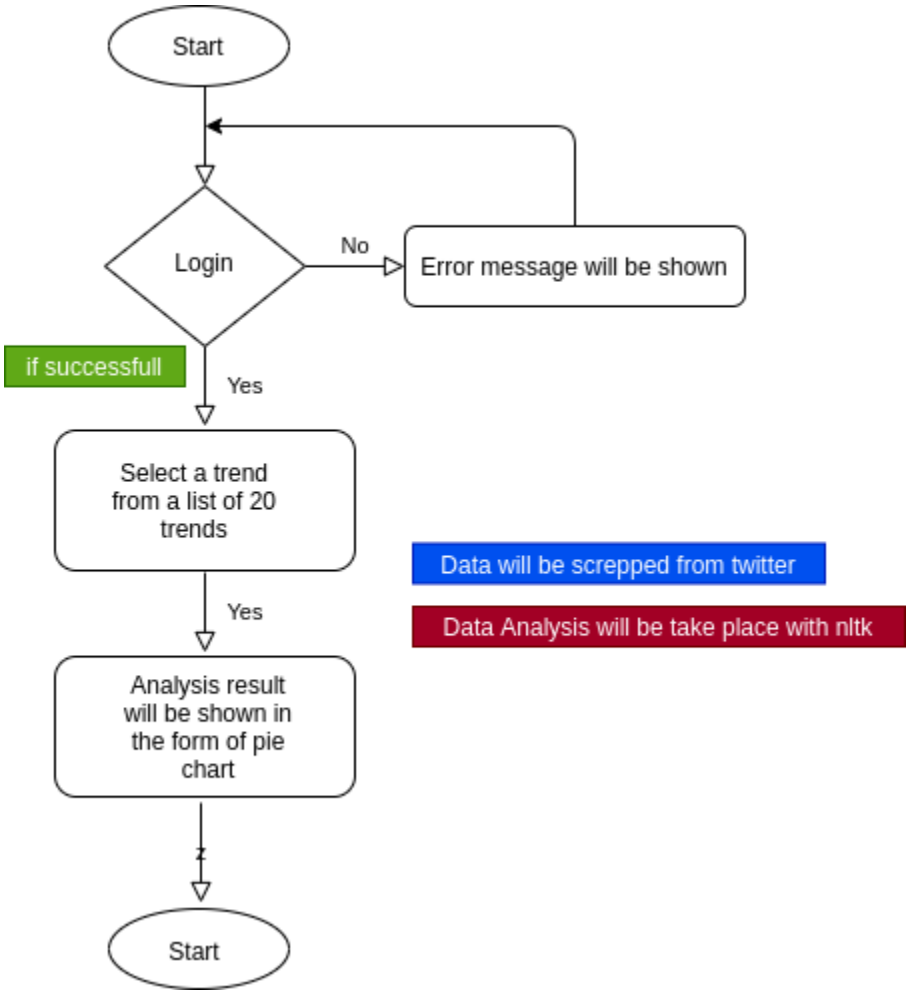
- Python 3 +
- Flask
- MongoDB
- Chrome driver (to scrape data from a twitter using selenium)
- Selenium (python package for web scrapping)
- Beautiful Soap (python package parse html pages)
- Numpy (python package for array manipulation with much optimization)
- Pandas (python package for tabular data manipulation)
- NLTK (python package natural language toolkit)
- Bootstrap 3+

After installing all the above package use following instructions

- Take a complete copy of the zip file from the given DVD and past it in your system
- And then open the project in any text editor
- Open terminal and write “python index.py” in windows or “python3 index.py” in Ubuntu.
- Then go to the address which is listed in the terminal and use the application
- Login page will be opened
- Enter credentials and click to login

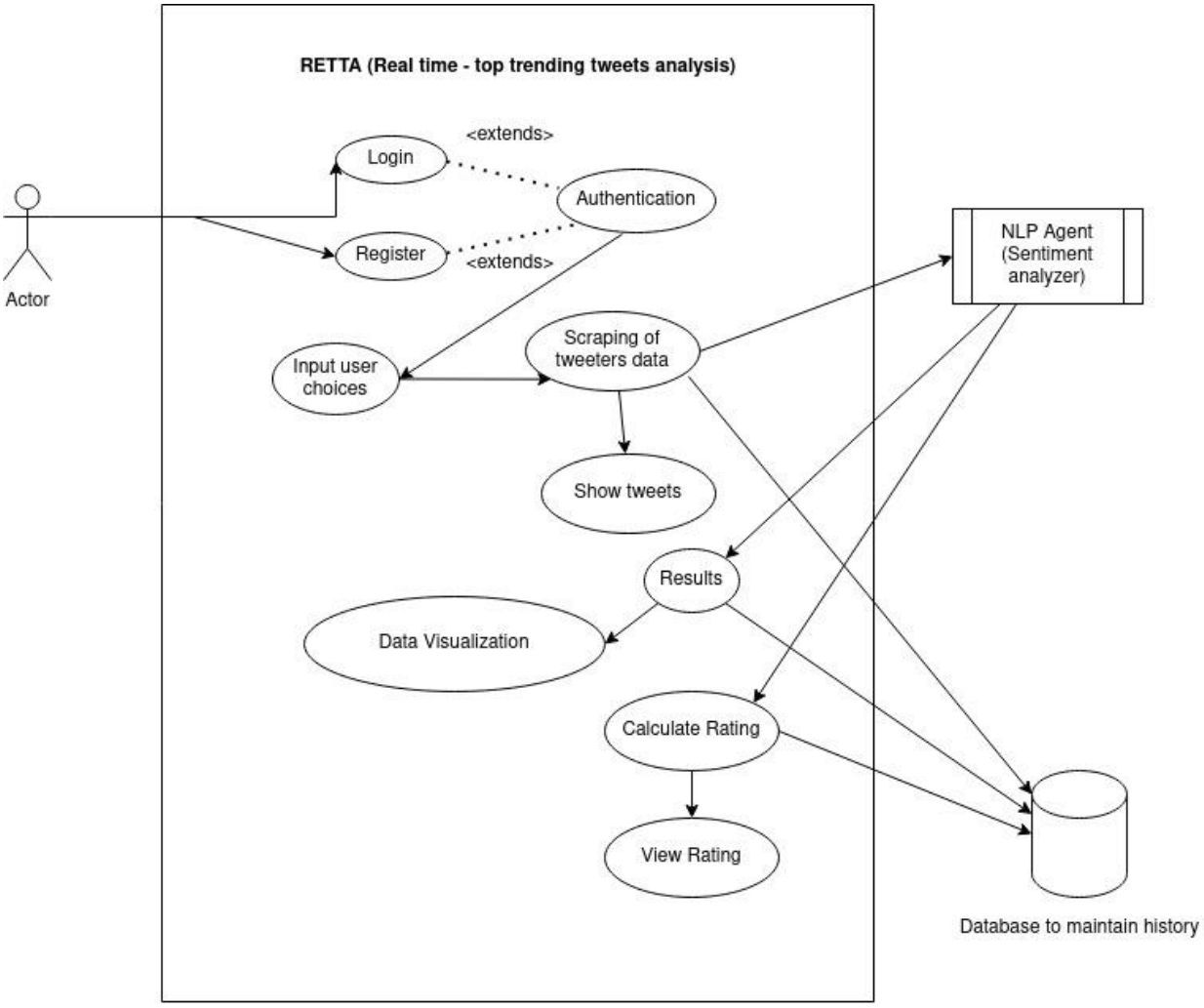
APPENDIX A

Flow Chart



APPENDIX B

Use-Case Diagram



APPENDIX C

Project Proposal

Extended Title: [If Required]

Real Time Tweet(top trending) Analysis (RETTA)

Brief Description of the Project / Thesis with Salient Space:

We look into the use of linguistic characteristics for identifying sentiment in trending tweets in this research. We use a supervised method to the problem, but we construct training data using existing hashtags from Twitter data. To cater for the problem and to reduce the time and effort in searching and analyzing the real time situation for a particular top trend we as a team have come with the proposal named as RETTA stands for "Real Time Tweet Analysis". As the name is self explinatory it highlights that our project will construct into a web application using latest machine learning techniques. It will provide user a service of Sentiment analysis of real time top trending tweets. RETTA will be different from other sentiment analysis web sites like Social Searcher, NCSU Tweet Sentiment Visualization App and Mention etc. RETTA will analyze only real time top trending tweets in Pakistan. None of the present website assist users with such type of functionality like real time tweets and most of such websites are paid.

Scope of Work:

- Scrapping data from Twitter for real time top trending tweets.
- Providing sentiment analysis of top trending tweets of Pakistan.
- Real time top trending tweets for sentiment analysis from a deployed machine learning model.
- Helping in knowing the masses opinion by sentiment analysis of real time top trending tweets.
- Keeping history of each previous request for future referencing.

Extended Scope:

1 Mobile Application

2 Real Time Data analysis & visualization

Academic Objective: This project would facilitate us (the team) in grasping in-depth knowledge besides getting hands-on experience on the latest methodologies and trends in the field of

Machine Learning, Scrapping, Data Science, Preprocessing, Web application development & integration

Application /End Goal Objective:

1 Providing analysis to digital marketing industry

2 Assists user in making right decisions

3 Minimizing search by Real time analysis

End Goals objective:

4 Real time top trending tweets for sentiment analysis from a deployed machine learning model

5 Front-end Application integration with deployed machine learning model / Application Programming Interface (API)

6 Sentiment Analyzer

7 End to end Framework

Previous Work Done on The Subject: No such service is available in Pakistan that helps you to do sentiment analysis of top trending tweets.

Material Resources Required: No major resources required to proceed with the project.

Data to be scrapped from Twitter.com

No of Students Required: 4

Special Skills Required:

2 Classical Machine Learning techniques

3 Python Language

4 Flask

5 Data Scrapping

6 Data Pre-processing

7 React.JS

REFERENCES

[1] NUST Regulations (Revised 2016): Part-II Academic Programs

[2] 2017, Software Requirements Specification for Academic Analysis System

[3] https://cms.nust.edu.pk

[4] 2014, Software Engineering Standards Committee of the IEEE Computer Society, "IEEE Recommended Practice for Software Design Descriptions", IEEE Std 1016-1998.

[5] Applying UML and Design Patterns - An Introduction to Object-Oriented Analysis and Design (Craig Larman) available from: http://www.ebookdirectory.com/

[6] UML basics: An introduction to the Unified Modeling Language (Donald Bell), available from: http://www.ebookdirectory.com/

[7] UML Distilled: A Brief Guide to the Standard Object Modeling Language by Martin Fowler, available from: http://www.ebookdirectory.com/

[8] Learning UML 2.0 by Russ Miles, available from: https://www.scribd.com/

[9] The Unified Modeling Language User Guide by Grady Booch, available from: https://www.scribd.com/

BIBLIOGRAPHY

[1] Text analyses Guide

[2] https://www.datacamp.com/community/tutorials/text-analytics-beginners-NLTK

[3] Sentiment analysis for exploratory data
https://programminghistorian.org/en/lessons/sentiment-analysis

[4] Web scraping using python

[5] https://www.analyticsvidhya.com/blog/2020/08/web-scraping-selenium-with-python/

[6] Web page parsing using beautiful soap

[7] https://www.dataquest.io/blog/web-scraping-python-using-beautiful-soup/

[8] Software Requirements (3rd Edition) (Developer Best Practices) 3rd Edition by Karl Wiegers