



**Department of Computer Science and
Engineering
Walchand College of Engineering, Sangli**

REPORT

T.Y in Computer Science and Engineering

Project Title

Sentimental Analysis Tool through Text-Mining

APPENDIX 1

Project Members

Name	Exam Seat No.	Roll No.	Email-Id	Contact
Vedant Sharma	2014BCS045	045	vedant.sharma200@gmail.com	9422252590
Anish Joshi	2014BCS049	049	joshianish18@gmail.com	9403068497
Mukund Sudharsan	2014BCS078	078	mukund.sudharsan@walchandsangli.ac.in	9920347210

Project Guide

Miss Padmashree.

Academic Year

2016-2017

APPENDIX 2

BONAFIDE CERTIFICATE

This is to certify that this project report entitled
“Sentimental Analysis Tool through Text-Mining”
submitted to **Walchand College of Engineering, Sangli** is a
bonafide record of work done by Vedant Sharma, Anish Joshi
and Mukund Sudharsan under my supervision from
“18/07/2016” to “05/11/2016”

Prof. Miss Padmashree

Dr. SMRITI BHANDARI

(Project Guide)

(HOD, Dept. Of C.S.E.)

Place: Sangli

Date: 17/11/2016

APPENDIX 3

Declaration by Authors

This is to declare that this report has been written by us. No part of the report is plagiarized from other sources. All information included from other sources have been duly acknowledged. We aver that if any part of the report is found to be plagiarized, we are shall take full responsibility for it.

Vedant Sharma 2014BCS045

Anish Joshi 2014BCS049

Mukund Sudharsan 2014BCS078

Place: Sangli

Date: 17/11/2016

APPENDIX 4

Table Of Content

Title	Page No.
Appendix 1	2
Appendix 2	3
Appendix 3	4
Table of Content	5
Technical Domain	6
Benefits of Project	7
Abstract	8
Project Overview	9
System Configuration	11
Objectives	11
Methodology	11
Object Oriented Design and Modelling Diagrams	13
Project Management	18
Bibliography	20

Technical Domain

Technical Area(s) Explored

- Data Mining.
- Text-Analysis.
- R-World.
- Naïve Bayes Classifier Algorithm.

Application Domain

- Sentimental Analysis.
- Business Intelligence.
- Linguistic description of texts.

Application Users

- Application to Review-Related Websites
 - Movie Reviewer, Product Reviewer, Poll Predictions.
- Application in Business and Government Intelligence
 - Knowing Consumer attitudes and trends.
- Application across Different Domains
 - Knowing public opinions for political leaders or their notions about rules and regulations in place etc.
- Application as a Sub-Component Technology
 - Detecting antagonistic ,heated language in mails,
 - Spam detection, context sensitive information detection etc.

Benefits of Project

- **Improve Customer Service**

Sentiment Analysis gives useful insights about your current and future customers' purchase preferences, brand affiliations, topics of interests, opinions, point of views on discussions, likes and dislikes in products/ services and much more. This useful information lets organizations to drastically improve their customer service and engagement strategies by building on the positive sentiments and formulating methods to combat negative sentiments.

- **Review Of Brands**

One of the best uses of Sentiment Analysis is that it allows organizations to quantify perceptions – about your brand, products and services, marketing campaigns, social engagement initiatives, online content etc. Organizations can use this information for devising better and more effective branding and marketing strategies and thus improve your brand reputation.

- **Beat Competition**

Sentiment Analysis lets organizations to know sentiments surrounding your competitors too. This allows you to benchmark your performance against that of your competitors. Using the reported sentiments, you can also predict trends and develop your specific social strategies to leverage these trends.

- **Gain Business Intelligence**

Sentiment Analysis empowers organizations by providing extensive, insightful information regarding their target audiences' sentiments. Made use of correctly, these sentiments are a gold mine of newer business possibilities and opportunities. Thus, Sentiment Analysis provides you insightful business intelligence using which you can take impactful decisions that would leverage your business.

- **Public Opinion**

Sentiment analysis is extremely useful in social media monitoring as it allows us to gain an overview of the wider public opinion behind certain topics.

Assumptions:

- User must have internet connection to access the web page.
- User must search for valid content.

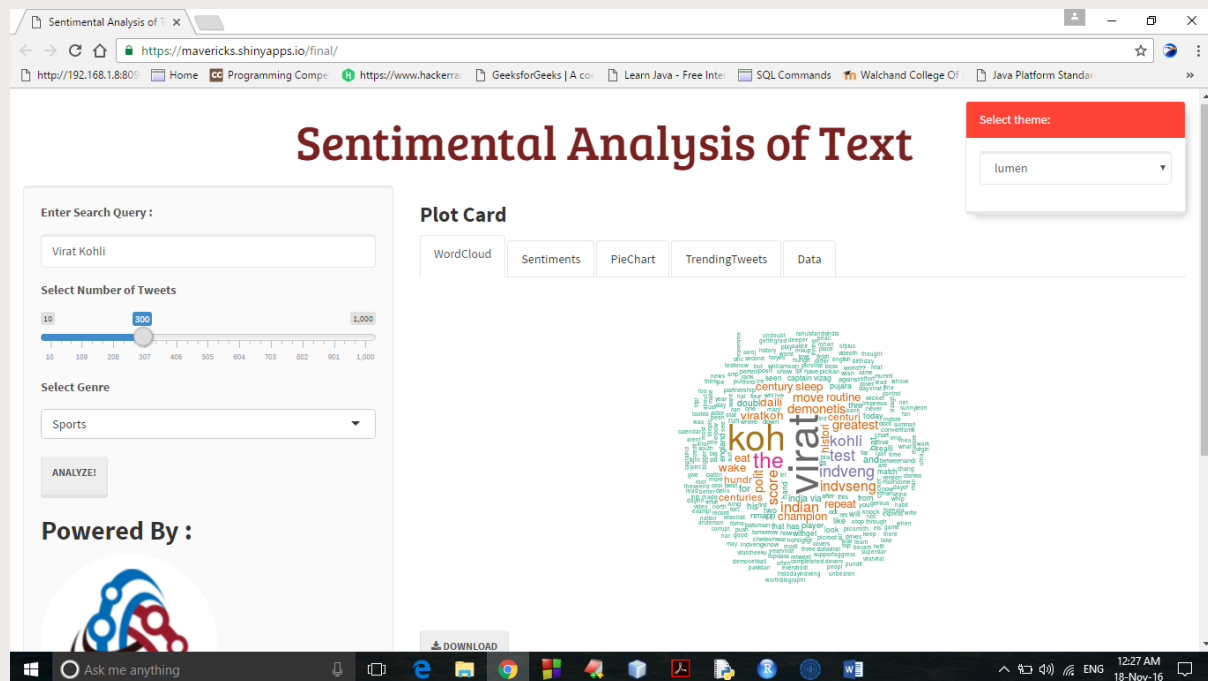
Abstract

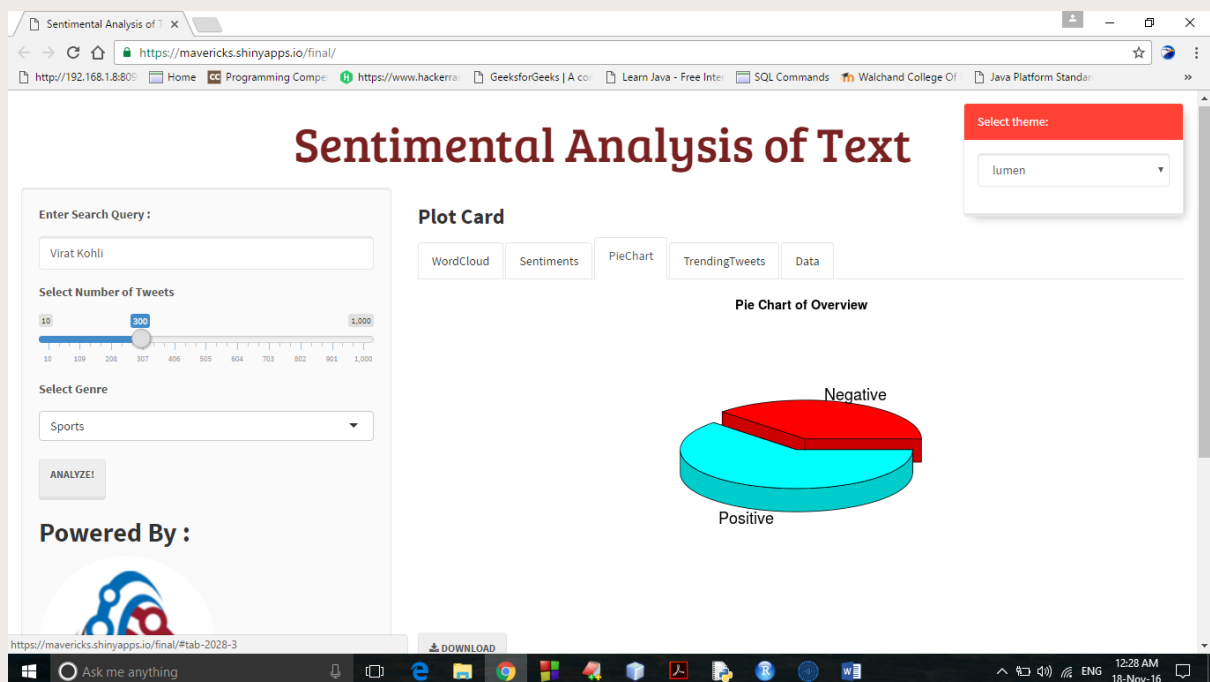
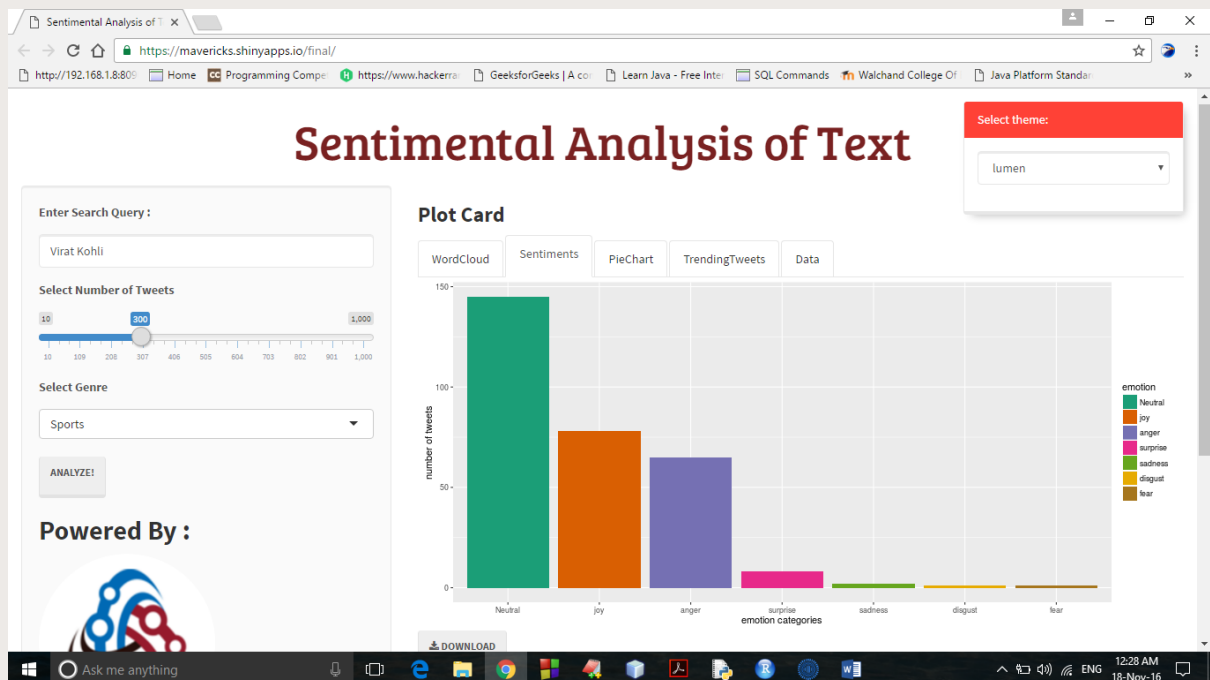
Sentiment analysis or opinion mining is the computational study of people's opinions, sentiments, attitudes, and emotions expressed in written language. Also it refers to the task of natural language processing to determine whether a piece of text contains some subjective information and what subjective information it expresses, i.e., whether the attitude behind this text is positive, negative or neutral. Understanding the opinions behind user-generated content automatically is of great help for commercial and political use, among others. The task can be

conducted on different levels, classifying the polarity of words or sentences.

It is one of the most active research areas in natural language processing and text mining in recent years. Its popularity is mainly due to two reasons. First, it has a wide range of applications because opinions are central to almost all human activities and are key influencers of our behaviours. Whenever we need to make a decision, we want to hear others' opinions. Second, it presents many challenging research problems, which had never been attempted before the year 2000. Part of the reason for the lack of study before was that there was little opinionated text in digital forms. It is thus no surprise that the inception and the rapid growth of the field coincide with those of the social media on the Web. In fact, the research has also spread outside of computer science to management sciences and social sciences due to its importance to business and society as a whole.

Project Overview:





System Configurations

Platform: RStudio and Shiny Apps.

Programming Languages: R

External Devices Used: None.

Standards Followed: R Coding Standards.

Objective

- To implement algorithm for automatic classification of text into positive, negative or neutral.
- Sentimental Analysis to determine the attitude of the mass is positive , negative towards the subject of interest.
- Graphical Representation of the sentiment in the form of various charts.

Methodology

There are 5 steps to analyse sentiment data and here's the graphical representation of the methodology to do the same.



Methods of Sentiment Analysis

- **Data Collection**

Consumers usually express their sentiments on public forums like the blogs, discussion boards, product reviews as well as on their private logs – Social network sites like Facebook and Twitter. Opinions and feelings are expressed in different way, with different vocabulary, context of writing, usage of short forms and slang, making the data huge and disorganized. Manual analysis of sentiment data is virtually impossible. Therefore, special programming languages like ‘R’ are used to process and analyse the data.

- **Text Preparation**

Text preparation is nothing but filtering the extracted data before analysis. It includes identifying and eliminating non-textual content and content that is irrelevant to the area of study from the data.

- **Sentiment Detection**

At this stage, each sentence of the review and opinion is examined for subjectivity. Sentences with subjective expressions are retained and that which conveys objective expressions are discarded. Sentiment analysis is done at different levels using common computational techniques like Unigrams, lemmas, negation and so on.

- **Sentiment Classification**

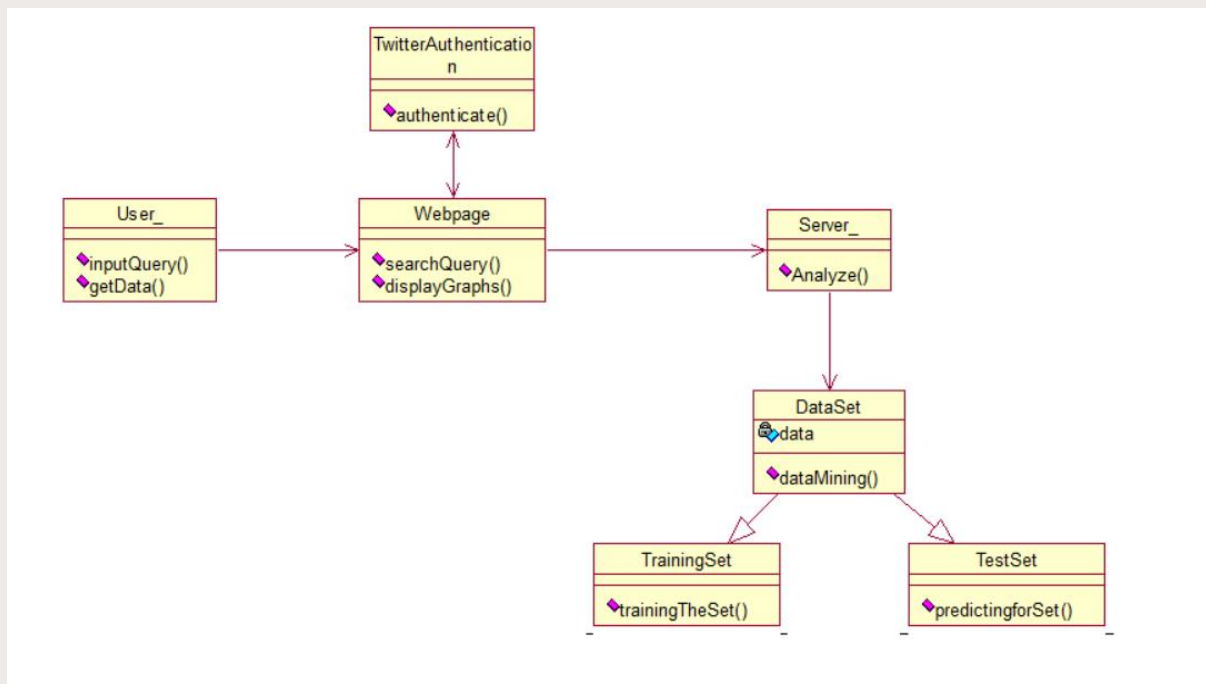
Sentiments can be broadly classified into two groups, positive and negative. At this stage of sentiment analysis methodology, each subjective sentence detected is classified into groups-positive, negative, good, bad, like, dislike.

- **Presentation of Output**

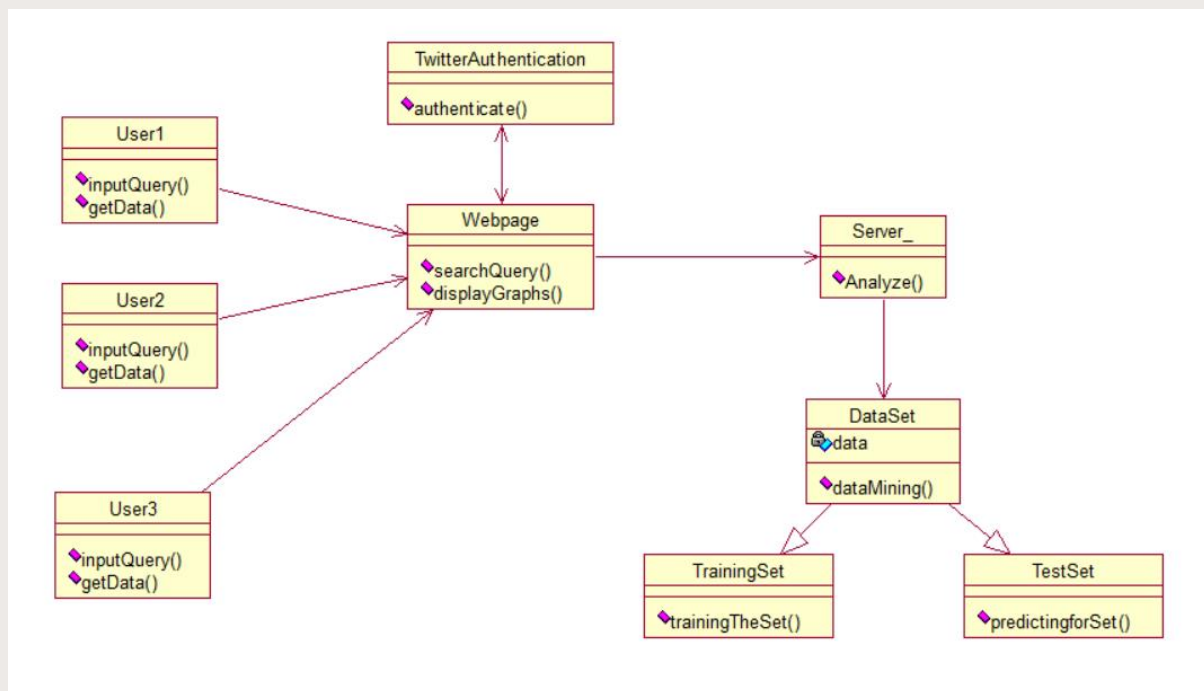
The main idea of sentiment analysis is to convert unstructured text into meaningful information. After the completion of analysis, the text results are displayed on graphs like pie chart, bar chart and line graphs.

Object Oriented Modelling and Design Diagrams:

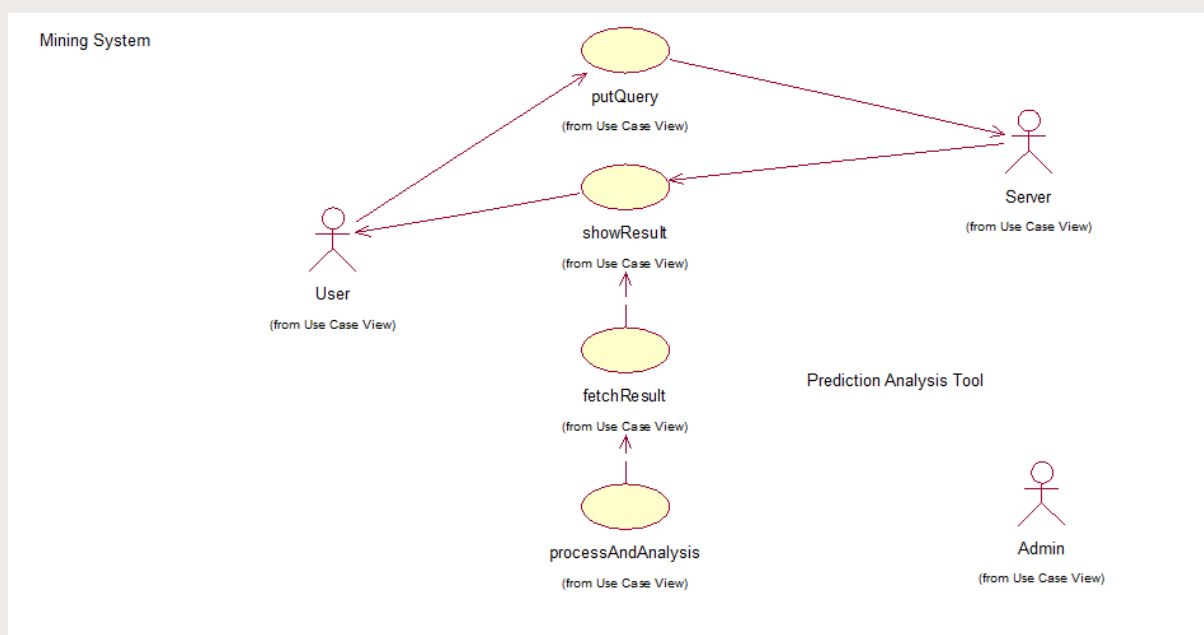
Class Diagram:



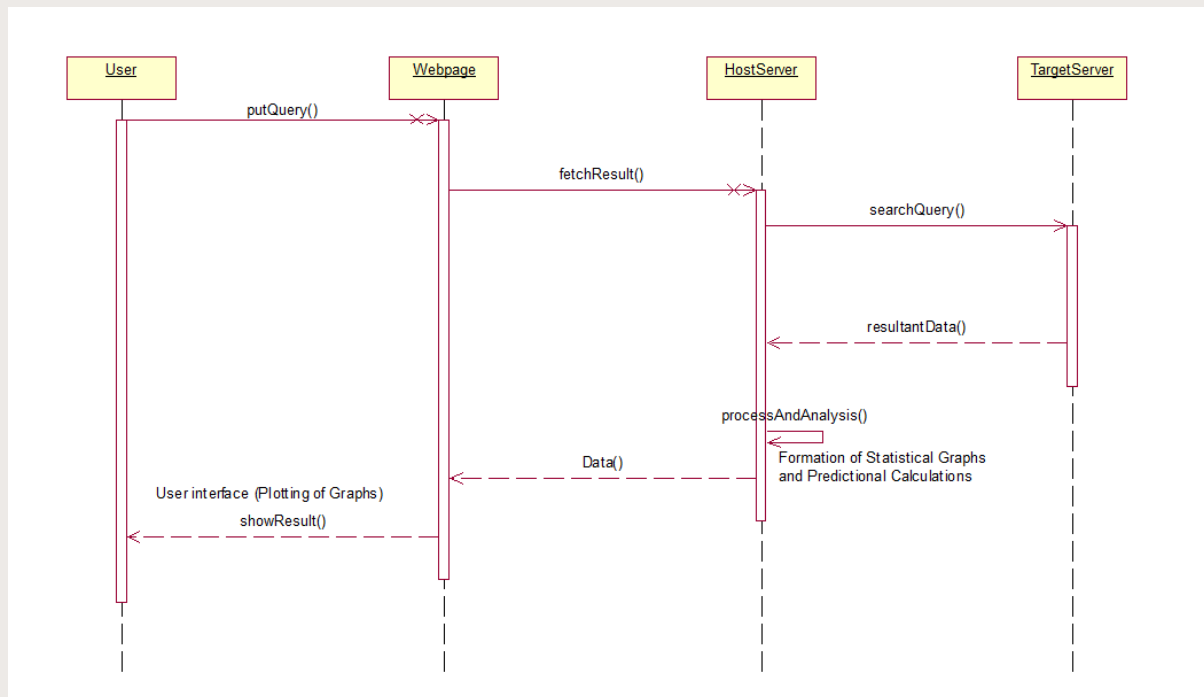
Object Diagram:



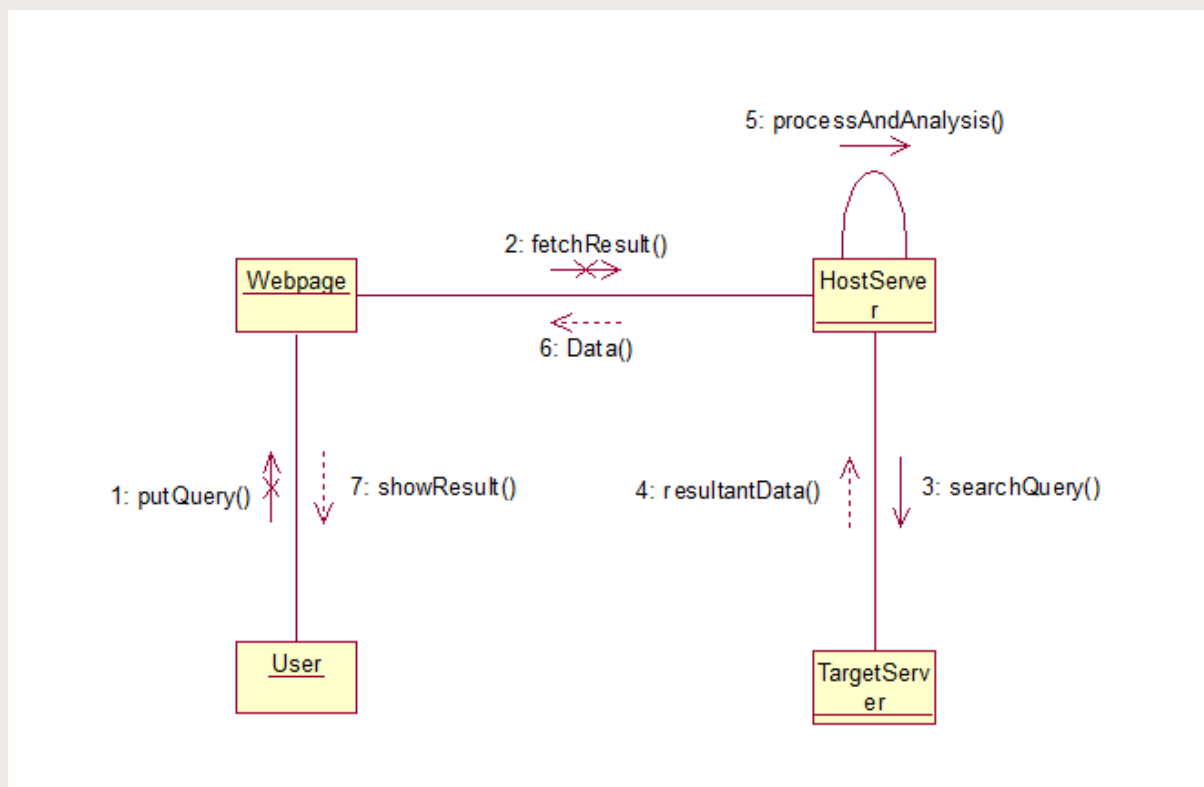
Use Case Diagram:



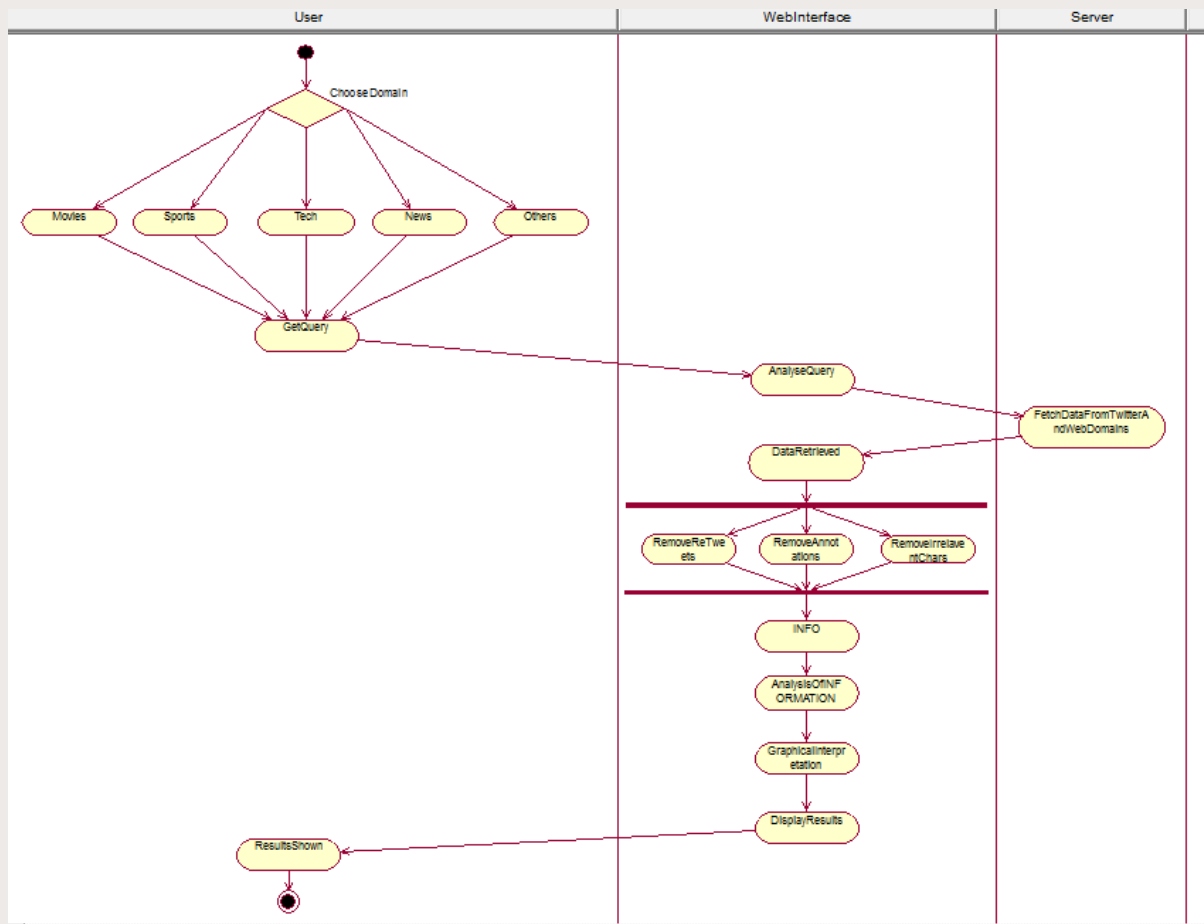
Sequence Diagram:



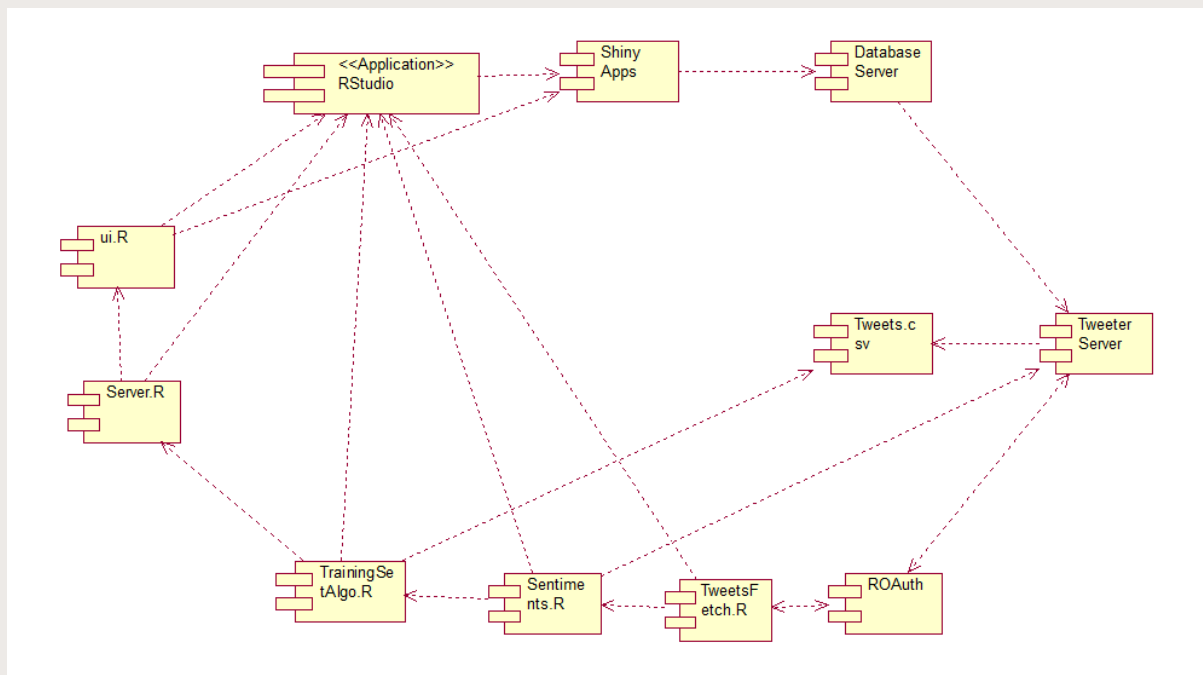
Collaboration Diagram:



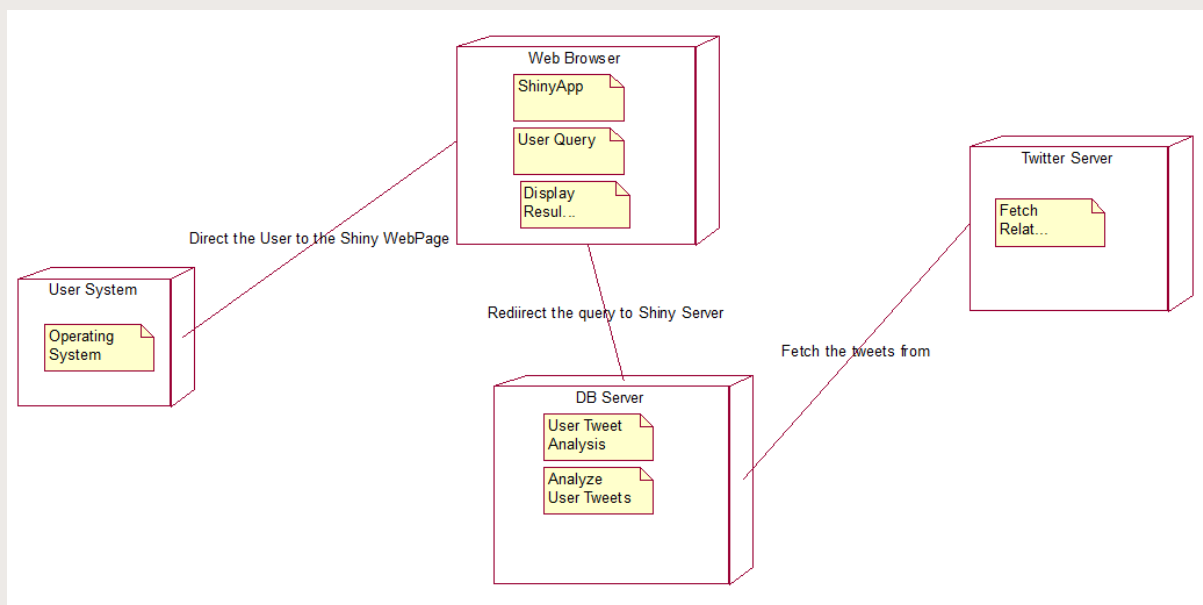
Activity Diagram:



Component Diagram:



Deployment Diagram:



PROJECT MANAGEMENT:

Gantt Chart Description:

Number	Milestone Name	Milestone Description	Timeline Number of weeks required to complete the milestone
1	Requirement Specification	A requirement specification document should be delivered.	1 week
2	Technology Familiarization	Understanding of technology. Each person should get themselves as expert in each of the technology and should arrange a half day session to share the info and come up with a document for reference	Working 3 week
3	System Setup	Setup up dev environment with the database servlet engine, also setup a test environment	1 week
4	Design	A high level architecture diagram and detailed design of all the modules. Also a datadictionary document should be delivered	2 week
5	Implementation	Naive Baye's Classifier algorithm is implemented	4 week
6	Debugging &	The code was	2 week

	Testing	tested for different kind of user inputs	
7	Deployment	The Webpage has been successfully deployed on shinyapps.io	3 days

Gantt Chart:

