ABSTRACT

Stress is essentially humans' response to various types of desires or threats. This response, when working properly,can help us to stay focused, energized and intellectually active, but if it is out of proportion, it can certainly be harmful leading to depression, anxiety, hypertension and a host of threatening disorders. Cyberspace is a huge soap box for people to post anything and everything that they experience in their day-to-day lives. Subsequently, it can be used as a very effective tool in determining the stress levels of an individual based on the posts and status updates shared by him/her. This is a proposal for a website which takes the Twitter username of the subject as an input, scans and analyses the subject's profile by performing Sentiment Analysis and gives out results. These results suggest the overall stress levels of the subject and give an overview of his/her mental and emotional state. The tool used for analysis of the social media account is Rapidminer. Rapid miner is an environment for various data mining and machine learning procedures with a very effective and simple GUI.[1]

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

1.1 Motivation

The successful implementation of an expert system depends very strongly on motivation. Along with the corporate goals of a company, motivation is the most vital factor. Before the start of expert systems one could obtain expert advice in two ways. In the first case, one could directly consult an expert. This process is expensive, if the expert is a professional person, such as a tax consultant; lawyer or a doctor. Furthermore, this means travelling to the expert and making an appointment which is not often easy. Sometimes people do not feel comfortable when discussing their personal problems with experts. There are exceptions, but this is the trend of public perception before people discuss their problem with an expert. In the second way, one could read the relevant books and articles, usually written by experts. This approach is not without its problems. It is time consuming to find the relevant books and articles. Moreover, it takes quite a long time to understand various concepts and find the particular part of the book which is relevant to one’s specific problem. Also, this approach is relatively suitable for a small part of the population which is really literate. Therefore, the most obvious reasons for building expert systems are, scarce human experts and the high cost of consulting them . The cost of developing an expert system is also high at the beginning, but in the long term it would be bearable compared with the growing cost of human expert employment. Human experts are not always within reach compared with expert systems. Furthermore, expert systems make it possible to computerize existing empirical knowledge within a company. This knowledge is available at anytime, anywhere, at constant quality. It is because of these reasons we felt that there is a need to develop a system which detects stress based on social interactions in social media.

1.2 Problem Definition

To create a website which will detect stress of any registered user based on his/her social media interactions. The website will scan the provided social media account and compute a stress score for the user’s account and any reactive measures, if required will be suggested.In our project, we use Twitter as case study, and explore the hypothesis of emotional contagion via the social stream. A reasonable expectation is that Twitter connections carry a small length of status than Facebook.

1.3 Relevance of the Project

Psychological stress is threatening people’s health. It is very important to detect stress timely for proactive care. Nowadays people can get all the information about another person via social networking sites and applications and user's mobile device contains all the necessary information about the user's current situations in the form of snaps or chats with his/her close friends.So we can say that stress level of the user is closely related to the text or media shared among friends.The rise of social media is changing people’s life, as well as research in healthcare and wellness. With the development of social networks like Twitter and Facebook, more and more people are willing to share their daily events and moods, and interact with friends through the social networks. As these social media data timely reflect users’ real-life states and emotions in a timely manner, it offers new opportunities for representing, measuring, modeling, and mining users behavior patterns through the large-scale social networks, and such social information can find its theoretical basis in psychology research.

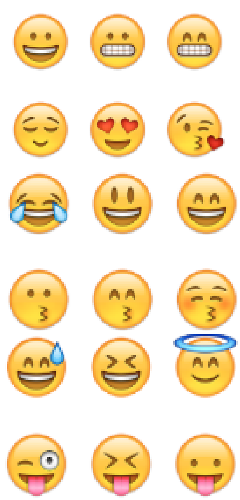
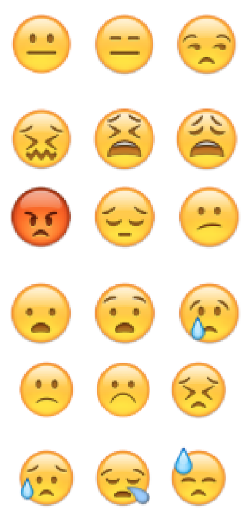
1.4 Methodology used

Twitter has over 328 million monthly active users. This makes up for a very large database to perform Natural Language Processing on. To study these tweets, we use an algorithm which is made up of 3 algorithms. These algorithms are:

1. The first part of the algorithm aims at retrieving the tweets which have certain keywords in them.
2. The second part of the algorithm performs Sentiment Analysis on the tweets found in the first step.
3. The third and final step calculates the resultant score based on the intensity values
4. assigned to the keywords.

The algorithm works as follows:

1. **Tweet retrieval:** There will be a dataset of keywords and emoticons which will be pre-saved into our system. This step will retrieve all those tweets which contain any one or more of these keywords and/or emoticons.
   The reason for including emoticons is that they have become a very famous means of dialogue and quite a lot of times, people just reply or communicate only using them.
   To carry out the process of retrieving the tweets from Twitter, a password, known as Twitter API key will be required which can be found from Twitter at request.
   Given below is a sample of these keywords and emoticons:

POSITIVE	**NEGATIVE**
Happy Excited Joyous Gleeful Satisfactory Funny Beaming	Unhappy Sorrowful Dejected Glum Gloomy Dismal Blue
POSITIVE	**NEGATIVE**
	

Keyword and Emoticon Table

2. **Sentiment Analysis:** The same words can be used in a variety of places to mean a lot of different things. We will need to determine what a particular keyword means in the context of the particular tweet or conversation.
   We will also be required to stem the words in order to determine the root words which will then be used for the purpose of analysing.
   The keywords will be segregated into three categories: positive,negative and neutral. Then, based on these scores, a final score will be computed which will be known as the complex result.

3. **Result declaration:** The complex result calculated in the above step will be the one which will be taken into account while declaring the final result about the stress of the particular user.
   There will be three categories in which these scores will be divided: unstressed, tensed and stressed. The highest scores will correspond to the unstressed category where the user, according to the tweets in analysis, is happy and satisfied. The tensed category will correspond to the moderate scores and it will mean the user is not stressed at the moment but proactive care needs to be taken. The stressed category will correspond to lowest scores and the users who will be in this category will be considered to tweet such messages or simply, tweets, which indicate that he is highly worried and stressed in his/her life.

CHAPTER 2: LITERATURE SURVEY

2.1. Research Paper

Paper 1: Detecting Stress Based on Social Interactions in Social Networks - Huijie Lin, Jia Jia*, Jiezhong Qiu, Yongfeng Zhang, Guangyao Shen, Lexing Xie, Jie Tang, Ling Feng, and Tat-Seng Chua

Abstract:

With the popularity people are used to sharing their daily activities and interacting with friends on social media platforms, making it feasible to leverage online social network data for stress detection. In this paper, we find that users stress state is closely related to that of his/her friends in social media, and we employ a large-scale dataset from real-world social platforms to systematically study the correlation of users' stress states and social interactions. We first define a set of stress-related textual, visual, and social attributes from various aspects, and then propose a novel hybrid model - a factor graph model combined with Convolutional Neural Network to leverage tweet content and social interaction information for stress detection.

Methodology Used:

The system includes the following phases:

- **Support Vector Machine (SVM):** It is a popular and binary classifier that is proved to be effective on a huge category of classification problems. In our problem we use SVM with RBF kernel.
- **Random Forest (RF):** it is an ensemble learning method for decision trees by building a set of decision trees with random subsets of attributes and bagging them for classification results.
- **Gradient Boosted Decision Tree (GBDT):** it trains a gradient boosted decision tree model with features associated with each user.
- **Deep Neural Network (DNN):** for user-level stress detection: it is proposed to deal with the problem of user-level stress detection problem with a convolutional neural network (CNN) with cross autoencoders. This is the real baseline method that we can compare our proposed model with.

Paper 2: Sentiment analysis in twitter using machine learning techniques-M. S. Neethu,R. Rajshri

Abstract:

Sentiment analysis deals with identifying and classifying opinions or sentiments expressed in source text. Social media is generating a vast amount of sentiment rich data in the form of tweets, status updates, blog posts etc. Sentiment analysis of this user generated data is very useful in knowing the opinion of the crowd. Twitter sentiment analysis is difficult compared to general sentiment analysis due to the presence of slang words and misspellings. The maximum limit of characters that are allowed in Twitter is 140. Knowledge base approach and Machine learning approach are the two strategies used for analyzing sentiments from the text. In this paper, we try to analyze the twitter posts about electronic products like mobiles, laptops etc using Machine Learning approach. By doing sentiment analysis in a specific domain, it is possible to identify the effect of domain information in sentiment classification. We present a new feature vector for classifying the tweets as positive, negative and extract people's' opinion about product.

Methodology Used:

Steps followed are as follows:

1.Symbolic Techniques:

In July 2013, Neethu M S and Rajasree R proposed that Symbolic techniques also known as knowledge based approach. In this technique, available lexical resources are used. In this sentiment analysis approach, bag-of-words approach is used. The BOW model focuses on the words list, or says string of words, it cannot check the context of the sentence. This model contains a list of words that have own value when found in the given text. This model totally focuses on the words and take care nothing about the language fundamentals.

2. Machine Learning Techniques:

In contrast to Knowledge based approaches, Machine Learning techniques are not using any lexicon resources list , instead a training set and a test set is used in order to classify them. Training set contains input vectors and corresponding class labels for training the network. After that, test set is used to validate the given model by checking the class labels to unknown feature vectors. There are different machine learning techniques like SVM, maximum entropy and Naïve Bayes etc. This allows the algorithm to remain dynamic in the face of ever changing social network language lexicons. In this methodology, a classification model is developed using a training set, which tries to classify the input feature vectors into corresponding class labels. Use the results from the knowledge based techniques and those of the machine learning techniques to ensure a thorough analysis of the dataset.

Paper 3: Detecting Emotions in Social Media: A Constrained Optimization Approach-Yichen Wang,Aditya Pal

Abstract:

Emotion detection can considerably enhance our understanding of users' emotional states. Understanding users' emotions especially in a real-time setting can be pivotal in improving user interactions and understanding their preferences. In this paper, we propose a constraint optimization framework to discover emotions from social media content of the users. Our framework employs several novel constraints such as emotion bindings, topic correlations, along with specialized features proposed by prior work and well-established emotion lexicons. We propose an efficient inference algorithm and report promising empirical results on three diverse datasets.

Methodology Used:

1.Sentiment Analysis:

Sentiment analysis aims at discovering the contextual polarity of the documents [Pang and Lee, 2008]. [Li et al., 2009] proposed a Non-negative Matrix Factorization (NMF) approach which leverages lexical knowledge for sentiment classification. Recent work [Bollen et al., 2011; Golder and Macy, 2011] has focused on mining temporal and seasonal trends of sentiment. Sentiment analysis is a closely related problem, however emotions are much more expressive than sentiments. Moreover, emotions need not contain a sentiment and vice-versa .

2.Emotion Detection:

Emotion models are primarily of two types [Ekkekakis, 2013]: (i) dimensional, and (ii) categorical. Dimensional models represent emotions on three dimensions: valence, arousal and dominance

3.Convex Sub-Problem

Paper 4: A Machine Learning Approach to Twitter User Classification

Abstract:

This paper addresses the task of user classification in social media, with an application to Twitter. We automatically infer the values of user attributes such as political orientation or ethnicity by leveraging observable information such as the user behavior, network structure and the linguistic content of the user's Twitter feed. We employ a machine learning approach which relies on a comprehensive set of features derived from such user information. We report encouraging experimental results on 3 tasks with different characteristics: political affiliation detection, ethnicity identification and detecting affinity for a particular business.

Methodology:

The system uses the following stages:

1.Profile features: "Who you are":

Most services (such as Twitter) publicly show by default profile information such as the user name, the location and a short bio. The Twitter API (2010) also provides access to other basic user information, such as the number of a user's friends, followers and tweets.

2.Tweeting behavior: "How you tweet":

Tweeting behavior is characterized by a set of statistics capturing the way the user interacts with the micro-blogging service: the average number of messages per day, number of replies, etc. Intuitively, such information is useful for constructing a model of the user; Java and colleagues (2007) suggest that users who rarely post tweets but have many followers tend to be information seekers, while users who often post URLs in their tweets are most likely information providers.

3.Linguistic content: "What you tweet":

Linguistic content information encapsulates the main topics of interest to the user as well as the user's lexical usage. Simple linguistic information is helpful for classifying users in several media, such as formal texts, blogs, spoken conversational transcripts or search sessions.

Paper 5: NLP based sentiment analysis on Twitter data using ensemble classifiers

Abstract:

Most sentiment analysis systems use bag-of-words approach for mining sentiments from the online reviews and social media data. Rather considering the whole sentence/ paragraph for analysis, the bag-of-words approach considers only individual words and their count as the feature vectors. This may mislead the classification algorithm especially when used for problems like sentiment classification. Traditional machine learning algorithms like Naive Bayes, Maximum Entropy, SVM etc. are widely used to solve the classification problems. These machine learning algorithms often suffer from biases towards a particular class. In this paper, we propose Natural Language (NLP) based approach to enhance the sentiment classification by adding semantics in feature vectors and thereby using ensemble methods for classification. Adding semantically similar words and context-sense identities to the feature vectors will increase the accuracy of prediction. Experiments conducted demonstrate that the semantics based feature vector with ensemble classifier outperforms the traditional bag-of-words approach with single machine learning classifier by 3-5%.

Methodology:

1. Ensemble methods for sentiment analysis:

In the field of ensemble methods, the main idea is to combine a set of models (base classifiers) in order to obtain a more accurate and reliable model in comparison with what a single model can achieve. The methods used for building upon an ensemble approach are many, and a categorization is presented in Rokach (2005). This classification is based on two main dimensions: how predictions are combined (rule based and meta learning), and how the learning process is done (concurrent and sequential).

2. Deep learning approaches:

In the realm of Natural Language Processing much of the work in deep learning has been oriented towards methods involving learning word vector representations using neural language models. Continuous representations of words as vectors has proven to be an effective technique in many NLP tasks, including sentiment analysis. In this sense, word2vec is one of the most popular approaches that allows modeling words as vectors. Word2vec is based on the Skip-gram and CBOW models to perform the computation of the distributed representations.

3. Ensemble taxonomy:

This section presents the proposed taxonomy for ensemble techniques applied to Sentiment Analysis in both surface and deep domains. This classification intends to summarize the work found in the literature as well as to compare these models with the ones we propose. Also, with this, we address the first question raised in Section 1 regarding how combination techniques can be classified.

Paper 6: Twitter sentiment analysis-Aliza Sarlan,Chayanit Nadam,Shuib Basri

Abstract:

Social media have received more attention nowadays. Public and private opinion about a wide variety of subjects are expressed and spread continually via numerous social media. Twitter is one of the social media that is gaining popularity. Twitter offers organizations a fast and effective way to analyze customers' perspectives toward the critical to success in the marketplace. Developing a program for sentiment analysis is an approach to be used to computationally measure customers' perceptions. This paper reports on the design of a sentiment analysis, extracting a vast amount of tweets. Prototyping is used in this development. Results classify customers' perspective via tweets into positive and negative, which is represented in a pie chart and html page. However, the program has planned to develop on a web application system, but due to limitation of Django which can be worked on a Linux server or LAMP, for further this approach need to be done.

Methodology:

Components of this system are:

1. Naive Bayes:

Naive Bayes is a conditional probability model that given a problem instance to be classified. Naive Bayes classifiers assume that the effect of a variable value on a given class is independent of the values of other variable.

2. SVM :

Vector machines (SVM) are primarily Classifiers that can classify by constructing hyperplanes that separate cases that belong to different categories. A Support Vector Machine (SVM) is a supervised classification algorithm that recently has been applied successfully to text classification tasks.

3. Maximum Entropy :

The Maximum entropy Classifier converts labeled feature sets to vectors using encoding. This encoded vector is then used to calculate weights for each feature that can then be combined to determine the most likely label for a feature set.

4. Random Forest:

Random Forest classifier is a tree-based classifier. It consists of numerous classification trees that can be used to predict the class label for a given data point based on the categorical dependent variable.

Paper 7: Twitter Sentiment Classification using Distant Supervision-Alec Go,Richa Bhayani,Lei Huang

Abstract:

We introduce a novel approach for automatically classifying the sentiment of Twitter messages. These messages are classified as either positive or negative with respect to a query term. This is useful for consumers who want to research the sentiment of products before purchase, or companies that want to monitor the public sentiment of their brands. There is no previous research on classifying sentiment of messages on microblogging services like Twitter. We present the results of machine learning algorithms for classifying the sentiment of Twitter messages using distant supervision. Our training data consists of Twitter messages with emoticons, which are used as noisy labels. This type of training data is abundantly available and can be obtained through automated means.

Methodology:

1. Defining Sentiment:
Many times it is unclear if a tweet contains a sentiment. For these cases, we use the following litmus test: If the tweet could ever appear as a frontpage newspaper headline or as a sentence in Wikipedia, then it belongs in the neutral class.

2 Characteristics of Tweets:
Length The maximum length of a Twitter message is 140 characters. From our training set, we calculate that the average length of a tweet is 14 words or 78 characters. This is very different from the previous sentiment classification research that focused on classifying longer bodies of work, such as movie reviews

3.Query Term:
We normalize the effect of query terms. Table 1 lists example query terms along with corresponding tweets. Our assumption is that users prefer to perform sentiment analysis about a product and not of a product.

4. Emoticons:
We strip the emoticons out from our training data. If we leave the emoticons in, there is a negative impact on the accuracies of the MaxEnt and SVM classifiers, but little effect on Naive Bayes. The difference lies in the mathematical models and feature weight selection of MaxEnt and SVM.

Paper 8: Sentiment analysis of twitter data using machine learning approaches and semantic analysis-

Abstract:
The wide spread of World Wide Web has brought a new way of expressing the sentiments of individuals. It is also a medium with a huge amount of information where users can view the opinion of other users that are classified into different sentiment classes and are increasingly growing as a key factor in decision making. This paper contributes to the sentiment analysis for customers' review classification which is helpful to analyze the information in the form of the number of tweets where opinions are highly unstructured and are either positive or negative, or somewhere in between of these two. For this we first pre-processed the dataset, after that extracted the adjective from the dataset that have some meaning which is called feature vector, then selected the feature vector list and thereafter applied machine learning based classification algorithms namely: Naive Bayes, Maximum entropy and SVM.

Methodology:

1.Random Walk : The random walk theory suggests that stock price changes have the same distribution and are independent of each other. So that the past movement or trend of a stock price cannot be used to predict its future movement.

2. Moving Average : It is widely used indicator in technical analysis that helps smooth out price action and filter out the noise from random price fluctuations. It is based on the past price of stock market. There are two commonly used MAs are the simple moving average (SMA), which is the simple average of a security over a defined number of time periods, and the exponential moving average (EMA), which gives bigger weight to more recent prices.

3. Regression method : Linear regression is most common technique for predicting the future value of variable based on linear relationship it has with other. It assume one straight line that approximates the given data, and forecast the future value based on direction of the regression line.

2.2 Books/Articles referred/ newspaper referred:

We have referred to the following websites to gather information related to our system:

1. "Introduction to Twitter Topic and Sentiment Analysis", Available: https://blog.algorithmia.com/anaylze-tweets-topic-sentiment-analysis/

   Inference drawn:
   The above website helps us in understanding how sentiment analysis on Twitter works. It gives an algorithm wherein a keyword is taken as an input from the user and the algorithm scans tweets and searches for the keyword among those tweets. Furthermore, the algorithm analyses the context in which the keyword is mentioned and gauges the reaction people have to that particular keyword. This is similar to our project in the sense that the above algorithm does Sentiment Analysis on Twitter and helps us understand how it works.

2. "What You Need to Know About Social Media Sentiment Analysis", Available: https://curatti.com/social-media-sentiment-analysis/

   Inference drawn:
   This website gives us an insight into how certain keywords actually affect the overall Sentiment Analysis for a particular topic or say, keyword. The inference which we draw from this website is that how sentiment analysis is actually done, what is the idea behind Sentiment Analysis and how it works. We are able to understand the process of Sentiment Analysis and the factors which affect its accuracy and which actually drive the entire process and all the ways in which this data curated can be used. It tells how Sentiment Analysis is powerful because it combines Human and Machine Learning.

CHAPTER 3. REQUIREMENTS

3.1 Functional Requirements:

Functional Requirements: Functional requirements may be calculations, technical details, data manipulation and processing and other specific functionality that define what a system is supposed to accomplish. Behavioral requirements describing all the cases where the system uses the functional requirements are captured in use cases.Generally, functional requirements are expressed in the form "system must do". The plan for implementing functional requirements is detailed in the system design.

3.1.1 Registration / Login Management:

- User will Login with their username and password.
- If user forgot password they can change the password by change password option.
- If user unable to login error message will be displayed.

3.1.2 Account Manager:

- User's different accounts will be accepted and on the basis of their comments and interactions, user will get the report.

3.1.3 Image/visual/text manager:

- This module includes management of text(chats,comments),images of the users shared on their account.

3.1.4 Stress Report:

- On the basis of user level attributes on the social interaction of the user.User must be given the stress level report.
- Additional to that user will be suggested do's and don'ts of that stress.

3.1.5 General Views :

- An user can click on any available link or page to view the content information.
- Users can exit the system at anytime.

3.1.6 Administrator Module:

- This module provides administrator related functionality.
- Administrator manages all information and has access rights to add, delete, edit and view the data related to user accounts, comments, interactions etc.

3.2 Non-Functional Requirements:

In systems engineering and requirements engineering, a non-functional requirement is a requirement that specifies criteria that can be used to judge the operation of a system,rather than specific behaviors. This should be contrasted with functional requirements that define specific behavior or functions.The plan for implementing non-functional requirements is detailed in the system architecture.non-functional requirements define how a system is supposed to be. Functional requirements are usually in the form of "system shall do ".

3.2.1 Reliability:

Stress detecting System shall be available 24 hours a day for application users.

3.2.2 Security:

Since the system needs access on the user's data which is in the form of text, visual,images etc. System shall not misuse or spread the data.

3.2.3 Performance:

Stress detecting System shall not take longer than 30 seconds to respond to a page request client; when using a standard internet connection.

3.2.4 Compatibility:

Stress detecting System application will be supported on current equipment such as desktop computers, laptops, printers, smartphones etc.

3.2.5 Interface:

Stress detecting System shall be accessible through a web browser such as Internet Explorer, Google Chrome or Mozilla Firefox. System shall provide printer friendly outputs of reports so that users can have easy to read printouts of the reports.

3.3 Constraints:

- The system provides solutions only based on Twitter posts and status.
- The system is only limited for the Twitter users.
- Internet connection is required for using this application.
- This system is not appropriate for high level stress detection
- The stress level shall be calculated only on the basis of the social media activity for the last 7 active days.

3.4 Hardware & Software Requirements:

Expert systems building tools (shells) are programs that make the job of building expert systems easier. Thus, tools are those programs which aid you to develop your own AI application. It has been claimed that these tools perform the development of expert systems in less time.

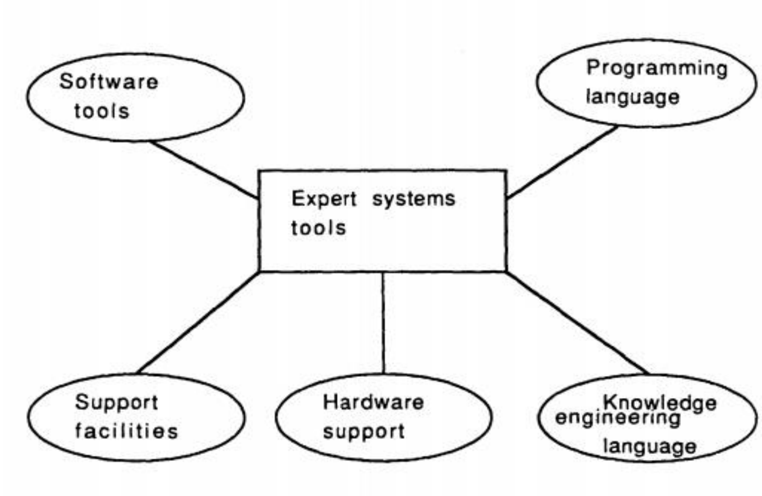


Figure 3.4.1 : Expert System Tools

Hardware Requirements:

- Processor - Pentium –IV
- RAM - 4 GB (min)
- Hard Disk - 20 GB
- Key Board - Standard Windows Keyboard
- Mouse - Two or Three Button Mouse
- Monitor - SVGA

Software Requirements:

- Operating System - Windows XP
- Coding Language - Java/J2EE(JSP,Servlet)
- Front End - J2EE
- Back End - MySQL

3.5 System Block Diagram

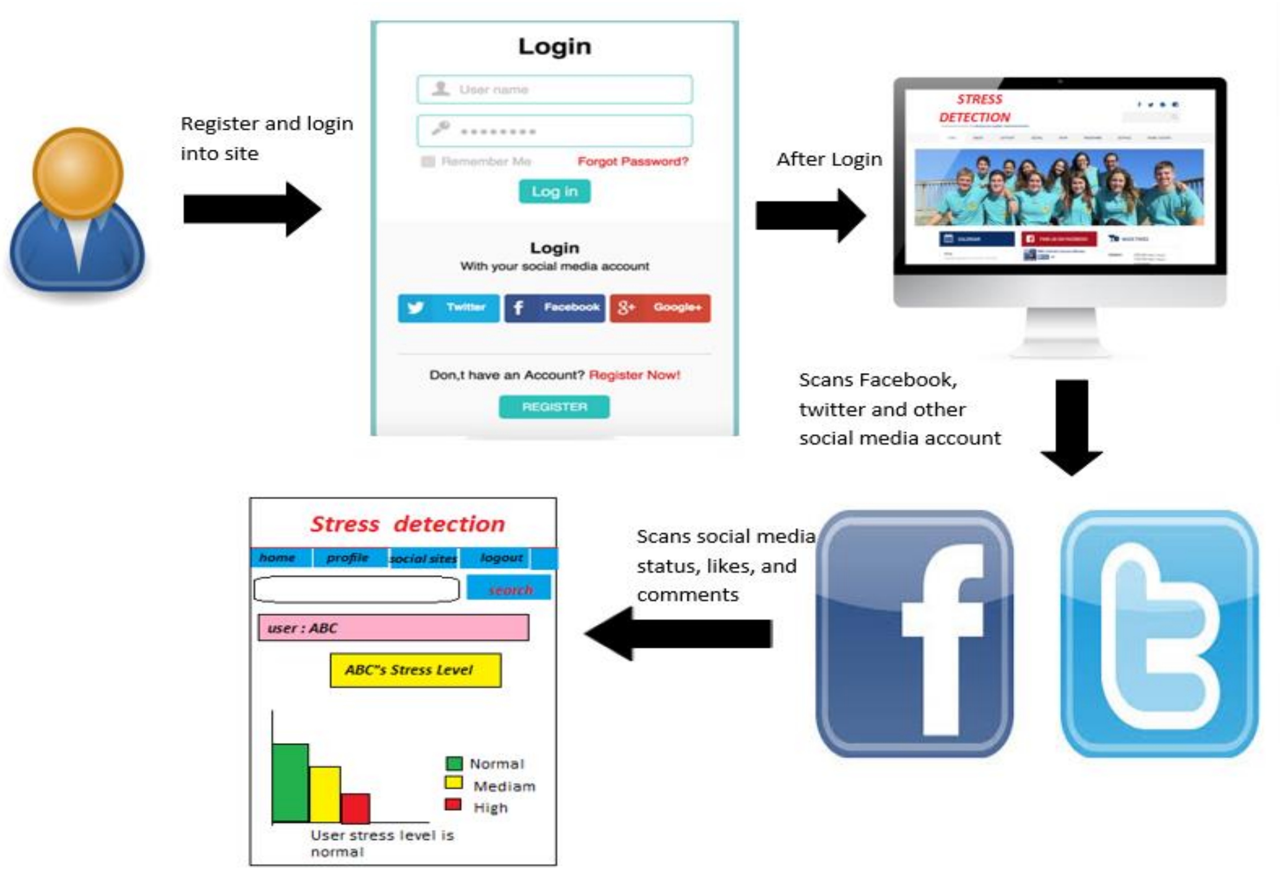


Fig 3.5.1: System block diagram for stress detection

This is the block diagram of "Detecting stress based on Social interaction in social networks".Here the user is interacting with our website through user interface which is made using html,css languages. First the user is login through our website and based on the user tweets stress is detected.There are total three categories whether the user's stress is positive,negative and neutral.If the user is stressed then doctor will provide advice to the user.

CHAPTER 4 : PROPOSED DESIGN

4.1 System Design / Conceptual Design (Architectural)

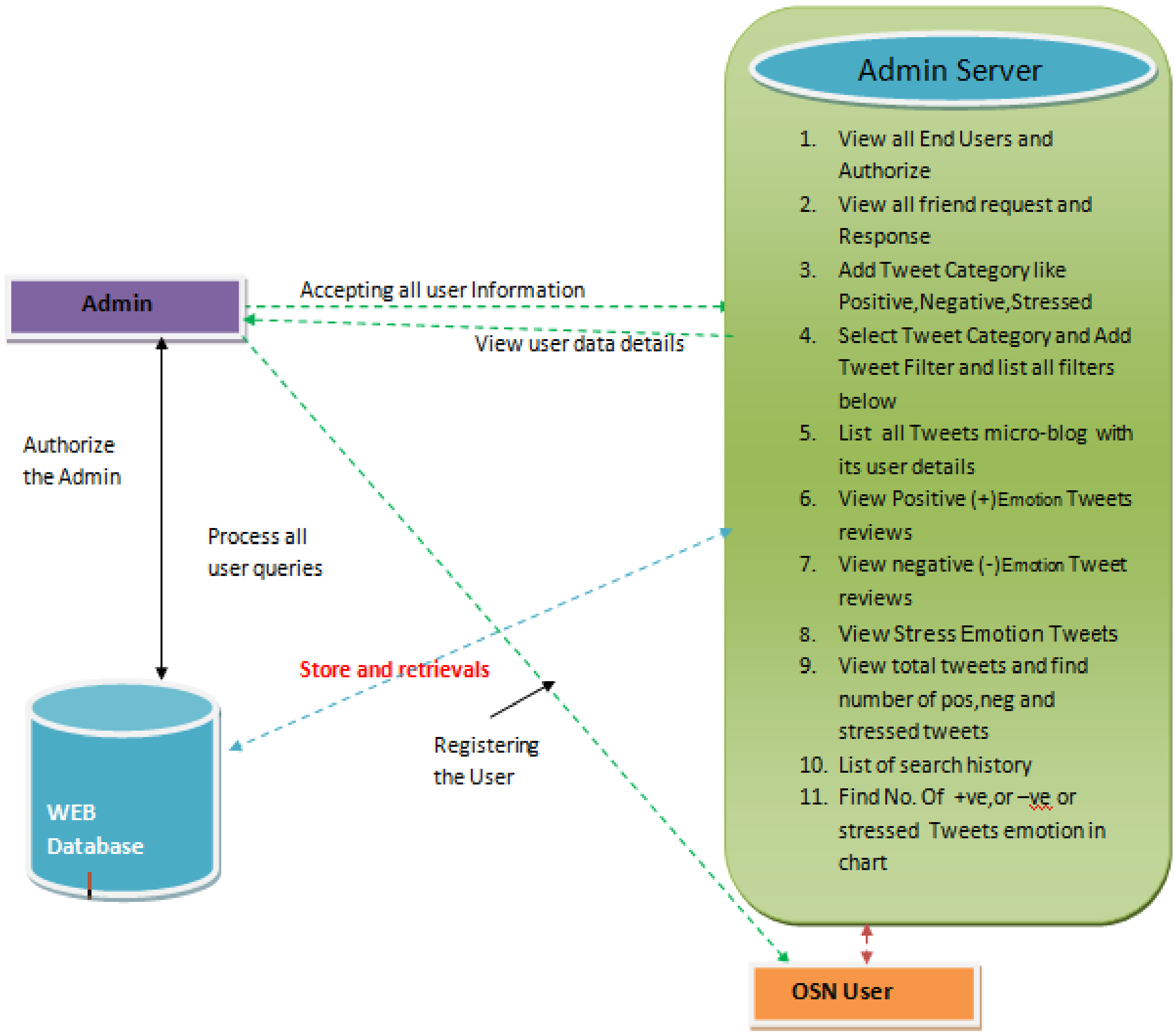


Fig 4.1.1: System Architectural diagram for stress detection

This is the architectural diagram of stress detection .There are two users OSN user and Admin user.Admin can view all the End users request and responses.OSN user can Register and login into the system .OSN user can search for the friends tweets and request friends,Create tweet by tweet name ,Tweet description ,Tweet image and date ,View all your created Tweets and find positive,negative,Stressed emotions on your Tweets.

4.2 Flowchart of the System

Flow chart for OSN User:

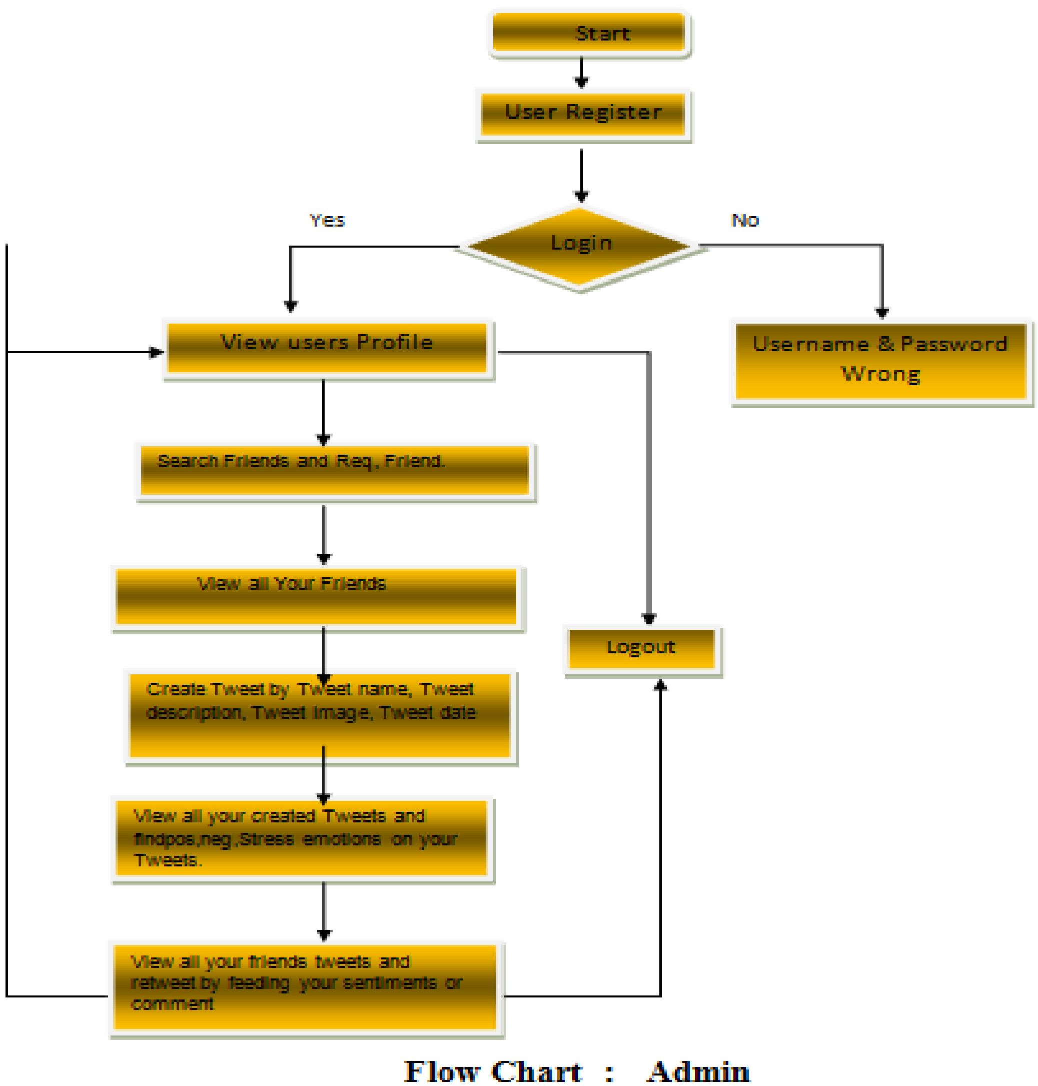


Fig 4.2.1: Flowchart for OSN User of stress detection

- **Flowchart for Admin User:**

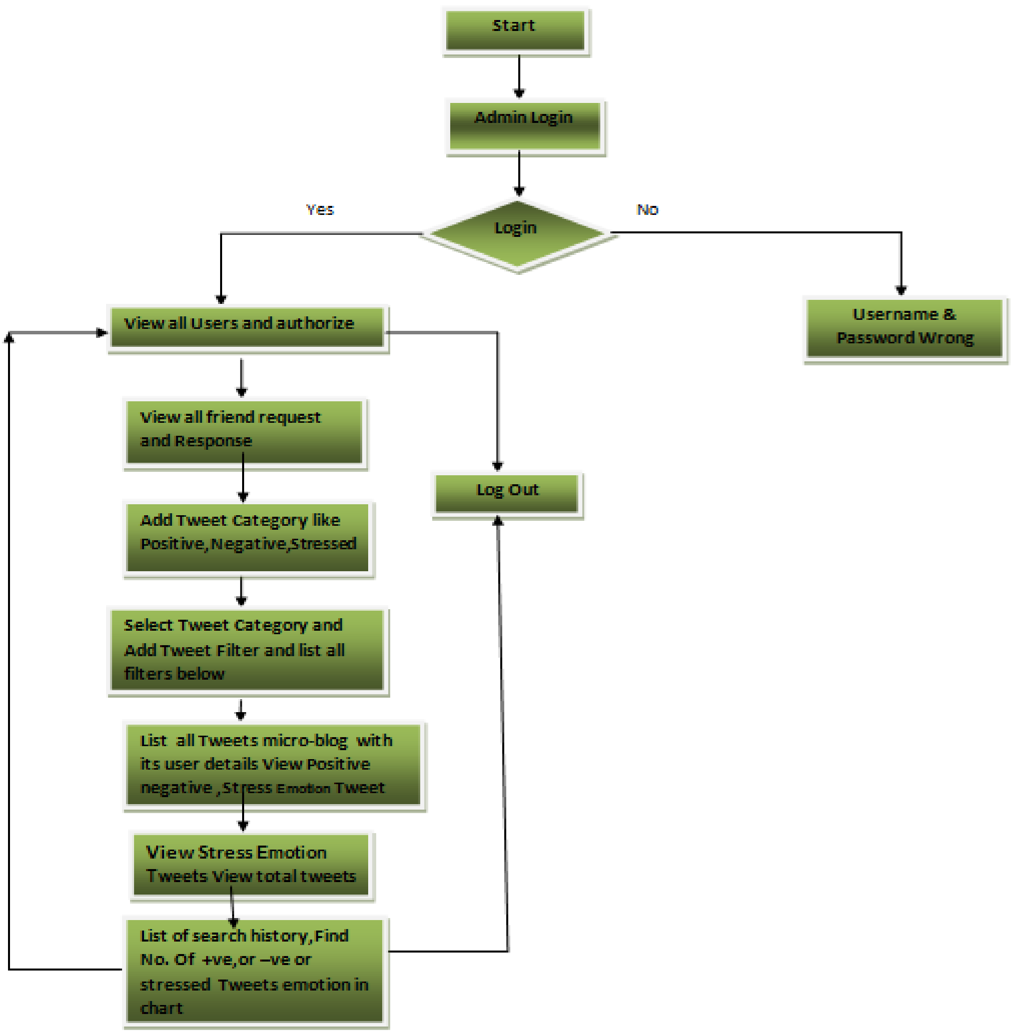
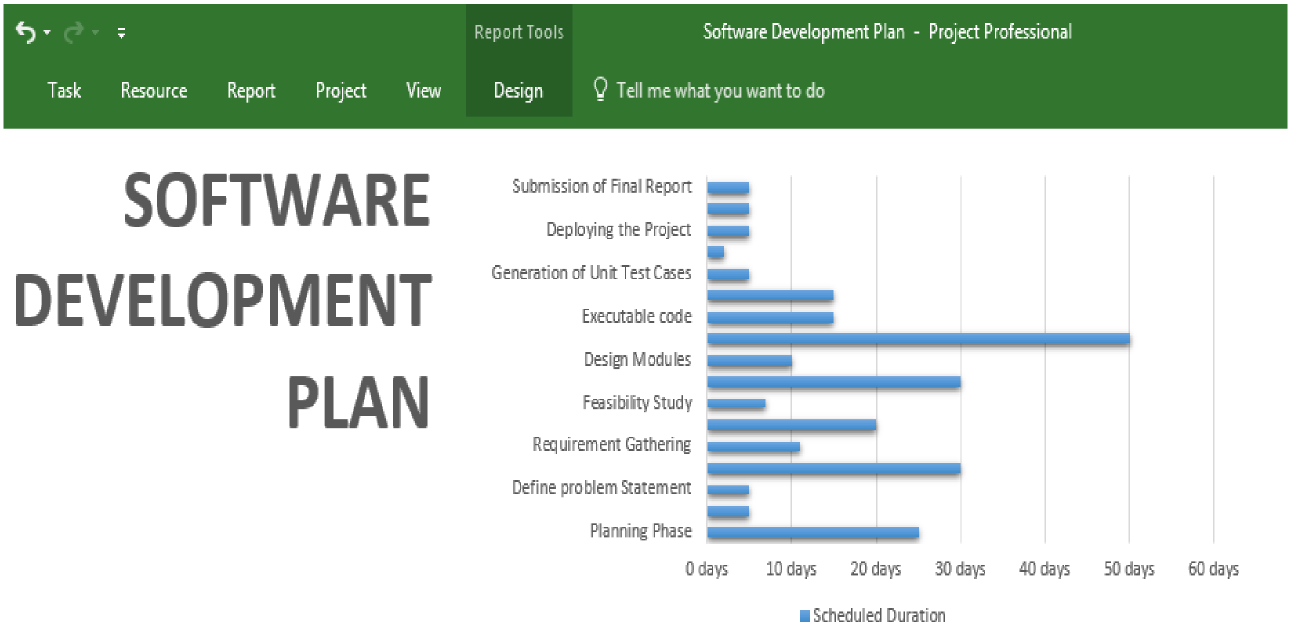


Fig 4.2.2: Flowchart for Admin User of stress detection

Flowchart shows a proper flow of activities of our application. Applications starts with Welcome screen. Now, if user has account then he/she can login if not he/she has to first sign up.User can search for friend ,request friend ,create tweet by tweet name ,And find the positive,negative and stressed tweets.Admin user can view all the users requests and add the tweet category like positive,negative ,stressed. Admin user can also view stress emotions.

4.3 Project Scheduling (Gantt Chart)

	Task Mode	Task Name	Duration	Start	Finish	Predecessors	Resource Names
0		**Software Development**	**422 days?**	**Thu 6/15/17**	**Fri 1/25/19**		
1		Planning Phase	25 days	Thu 6/15/17	Wed 7/19/17		Al,An,Sk,LB
2		Users Need	5 days	Thu 7/20/17	Wed 7/26/17	1	Al,An,Sk,LB
3		Define problem Statement	5 days	Thu 7/27/17	Wed 8/2/17	2,1	Al,An,Sk,LB
4		Literature Survey and Interviews	30 days?	Thu 8/3/17	Fri 3/23/18	2,3	Al
5		Requirement Gathering	11 days	Mon 3/26/18	Wed 10/3/18	4	Al,An,Sk,LB
6		Risk Analysis	20 days	Mon 3/26/18	Fri 4/20/18	4	LB
7		Feasibility Study	7 days?	Thu 10/4/18	Fri 10/12/18	5,6	Sk
8		Engineering Phase	30 days?	Thu 10/4/18	Wed 11/14/18	1,5	Al,An,Sk,LB
9		Design Modules	10 days	Thu 10/4/18	Wed 10/17/18	5	An
10		Genaration of Source Code	50 days	Thu 10/4/18	Wed 12/12/18	5,6	Al,An,Sk,LB
11		Executable code	15 days	Thu 12/13/18	Wed 1/2/19	10	Al,An,Sk,LB
12		Evaluation Phase	15 days	Thu 1/3/19	Wed 1/23/19	11	Al
13		Generation of Unit Test Cases	5 days	Thu 1/3/19	Wed 1/9/19	11	LB,Sk
14		Unit Test Cases Results	2 days	Thu 1/10/19	Fri 1/11/19	13	Al,An,Sk,LB
15		Deploying the Project	5 days?	Mon 1/14/19	Fri 1/18/19	11,14	Al
16		User Manual and Installation Guide	5 days?	Mon 1/21/19	Fri 1/25/19	2,15	An,Sk,LB
17		Submission of Final Report	5 days?	Tue 3/27/18	Mon 4/2/18		Al,An,Sk,LB



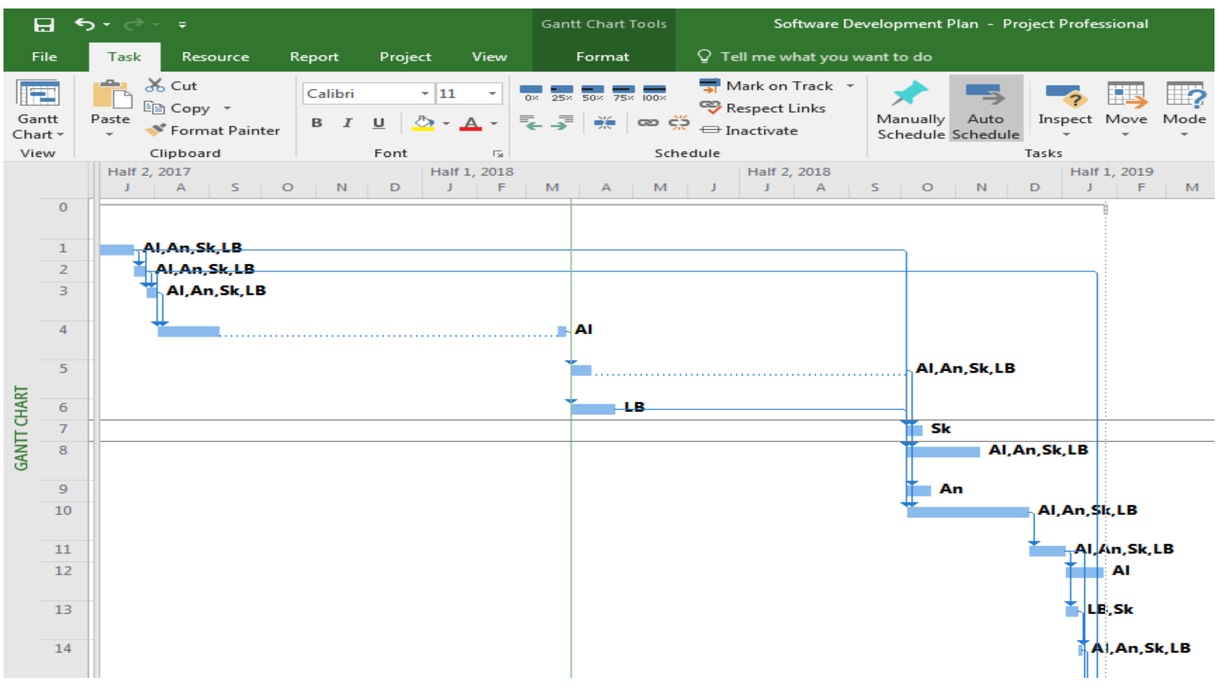
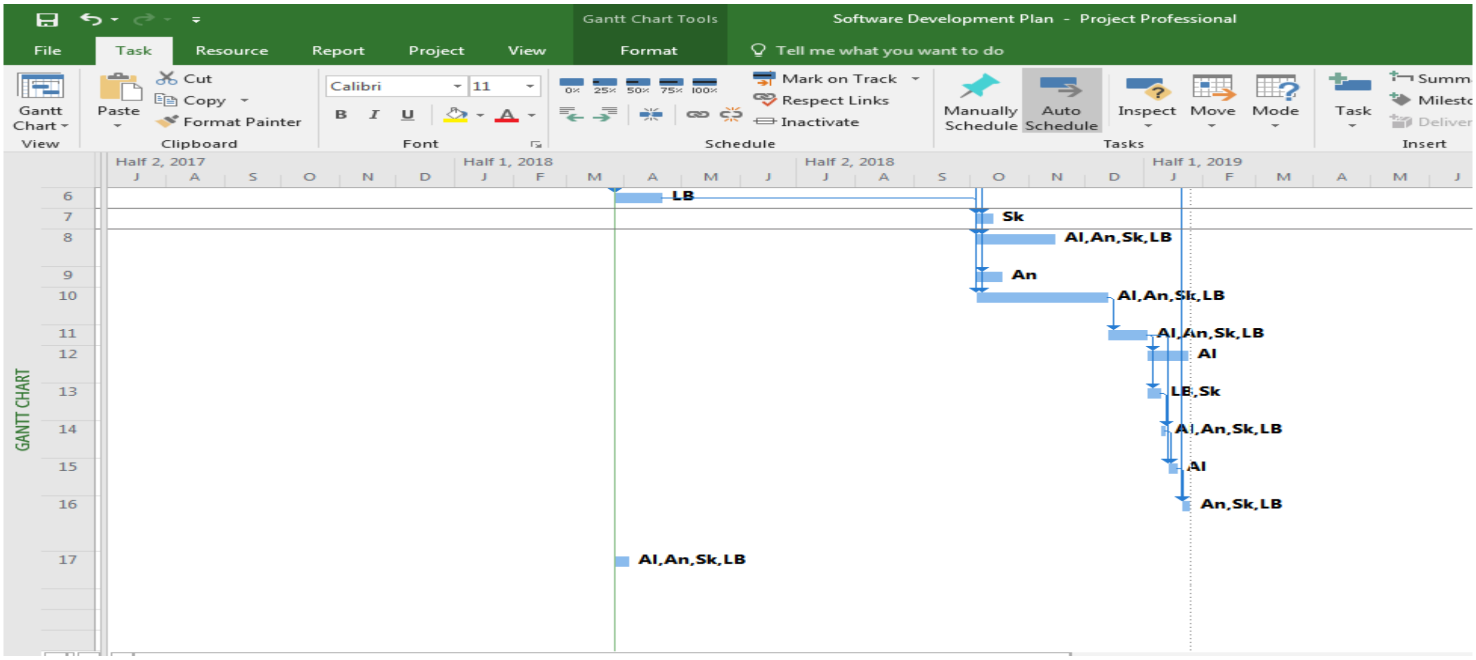
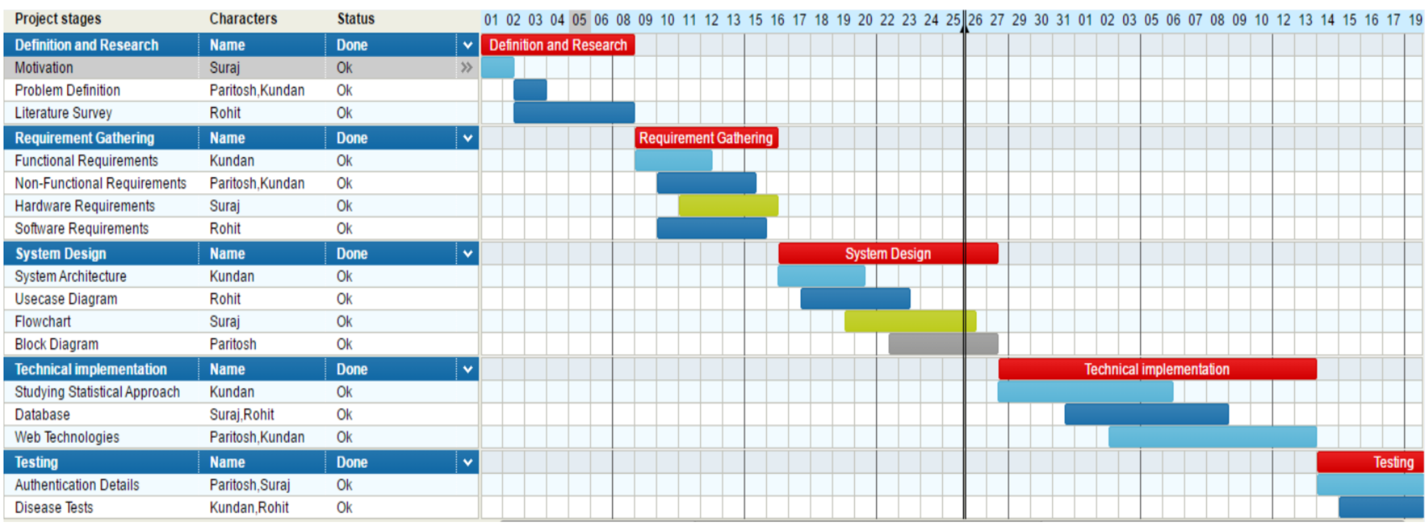


Figure 4.3.1: Task in Gantt Chart and Timeline





Project stages	Characters	Status
Definition and Research	Name	Done
Motivation	Suraj	Ok
Problem Definition	Paritosh,Kundan	Ok
Literature Survey	Rohit	Ok
Requirement Gathering	Name	Done
Functional Requirements	Kundan	Ok
Non-Functional Requirements	Paritosh,Kundan	Ok
Hardware Requirements	Suraj	Ok
Software Requirements	Rohit	Ok
System Design	Name	Done
System Architecture	Kundan	Ok
Usecase Diagram	Rohit	Ok
Flowchart	Suraj	Ok
Block Diagram	Paritosh	Ok
Technical implementation	Name	Done
Studying Statistical Approach	Kundan	Ok
Database	Suraj,Rohit	Ok
Web Technologies	Paritosh,Kundan	Ok
Testing	Name	Done
Authentication Details	Paritosh,Suraj	Ok
Disease Tests	Kundan,Rohit	Ok

The Gantt Chart of a system represents activities (tasks or events) displayed against time. It is drawn to represent a schedule of the whole project and also to show the dependency relationships between activities and current schedule status. First diagram represents a list of the activities and along the top is a suitable time scale. Each activity is represented by a bar; the position and length of the bar reflects the start date, duration and end date of the activity.

4.4 Entity Relationship Diagram

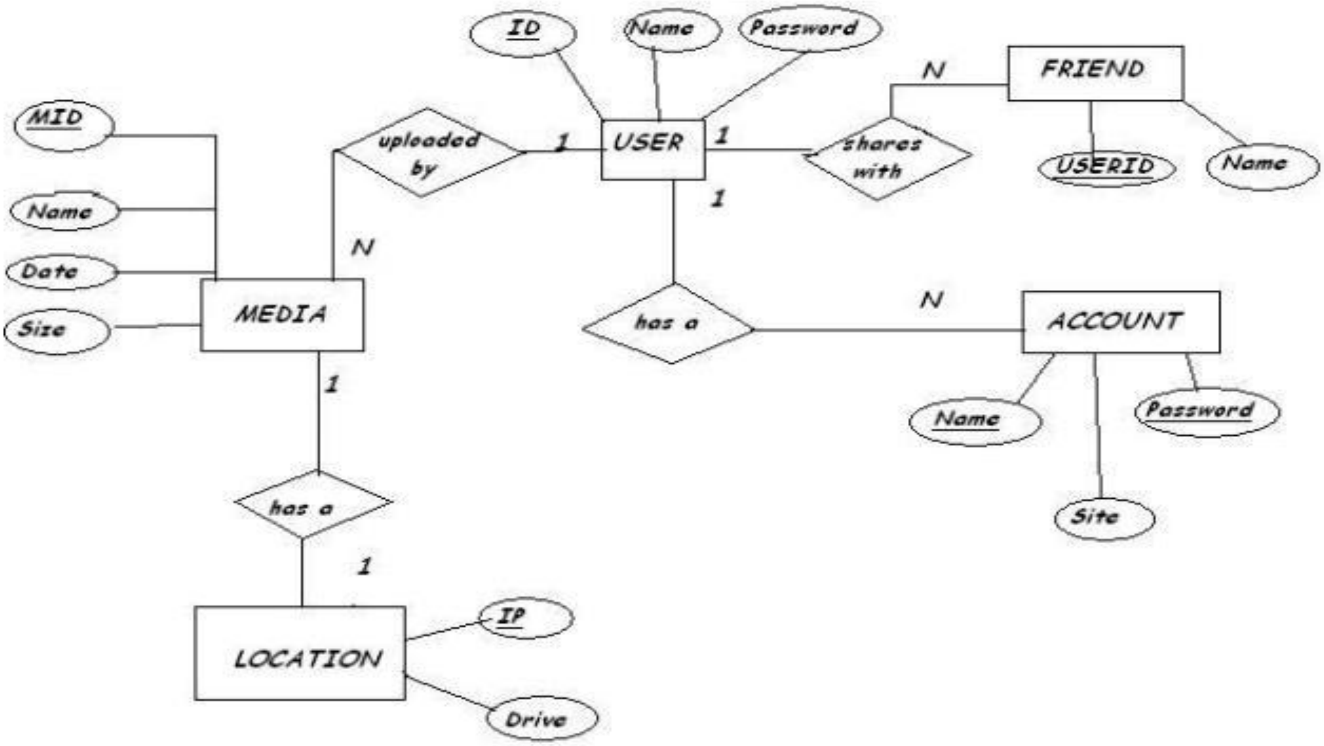


Fig 4.4.1: ER diagram for stress detection

This is Entity Relationship Diagram of Detecting stress based on social interactions in social networks. It shows the relationship between entities. There are in all four entities namely User, Media,Location and Account.

4.5 Data Flow Diagram:

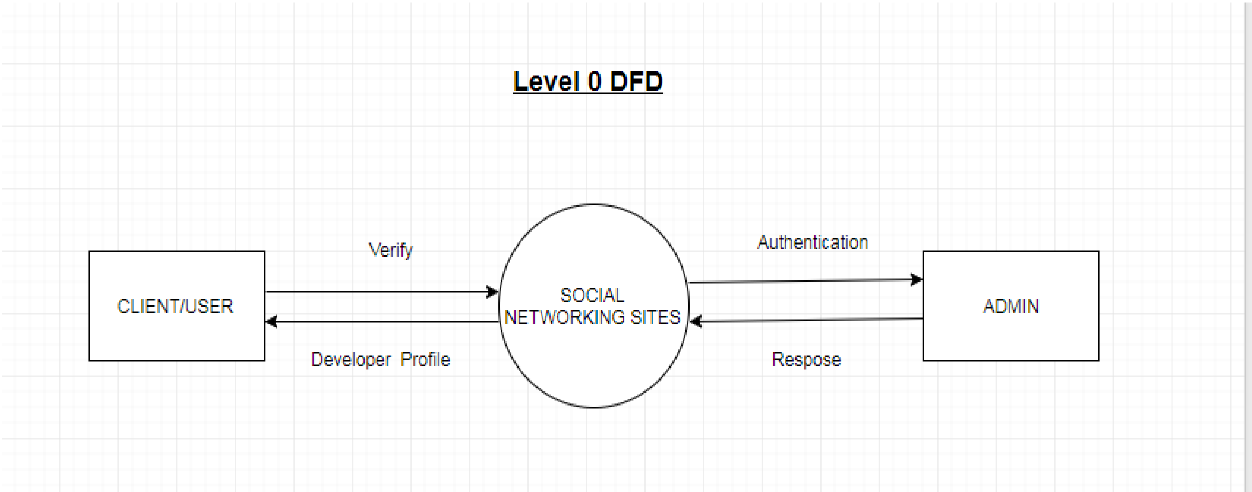


Figure 4.5.1: DFD Level 0

This is the Level 0 DFD Diagram. Our System has two primary modules. One is the User and other is the Admin. In this level, the system user will login into the website and the admin will provide the response to the user through system.

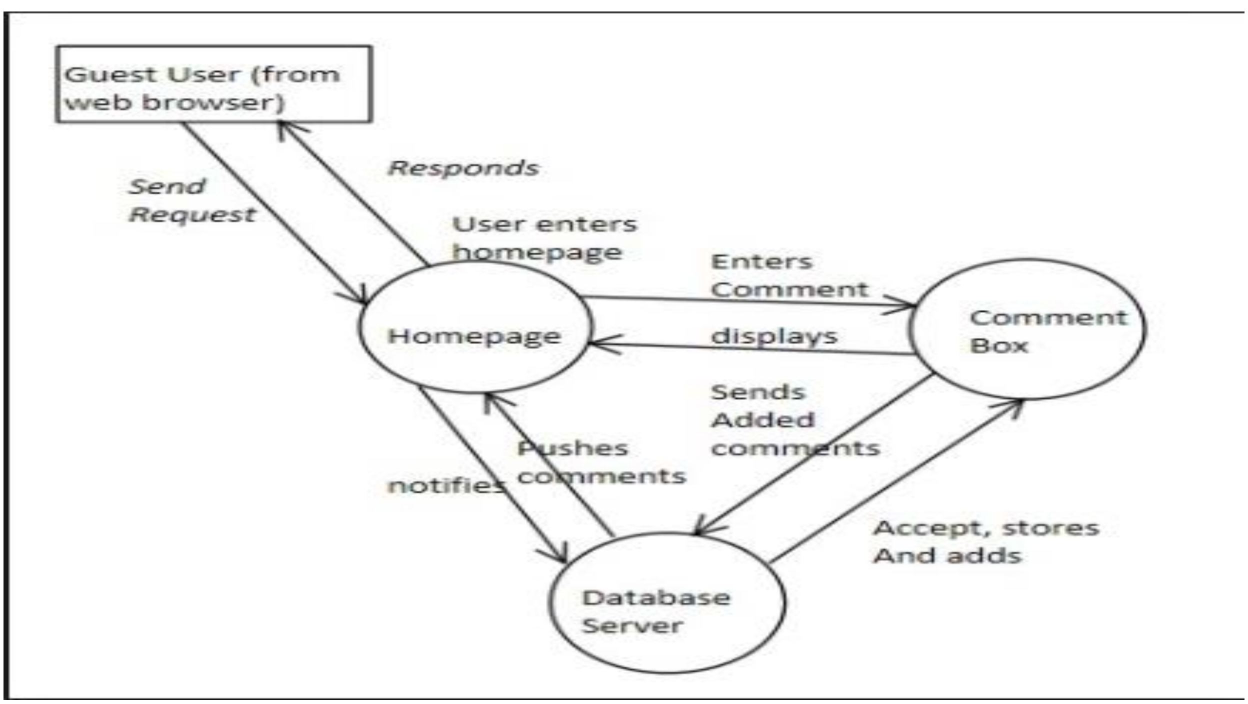


Figure 4.5.2: DFD Level 1

This is the DFD level 1 Diagram. It gives a flow of our system in a more detailed manner. There are several stages. In the first stage the user registers. If the user holds an account, then he/she can go for further steps such as the User will tweet on the social networking website and it will display on the homepage. Once the analysis is done, the system will then generate the results and it is then forwarded to the user.

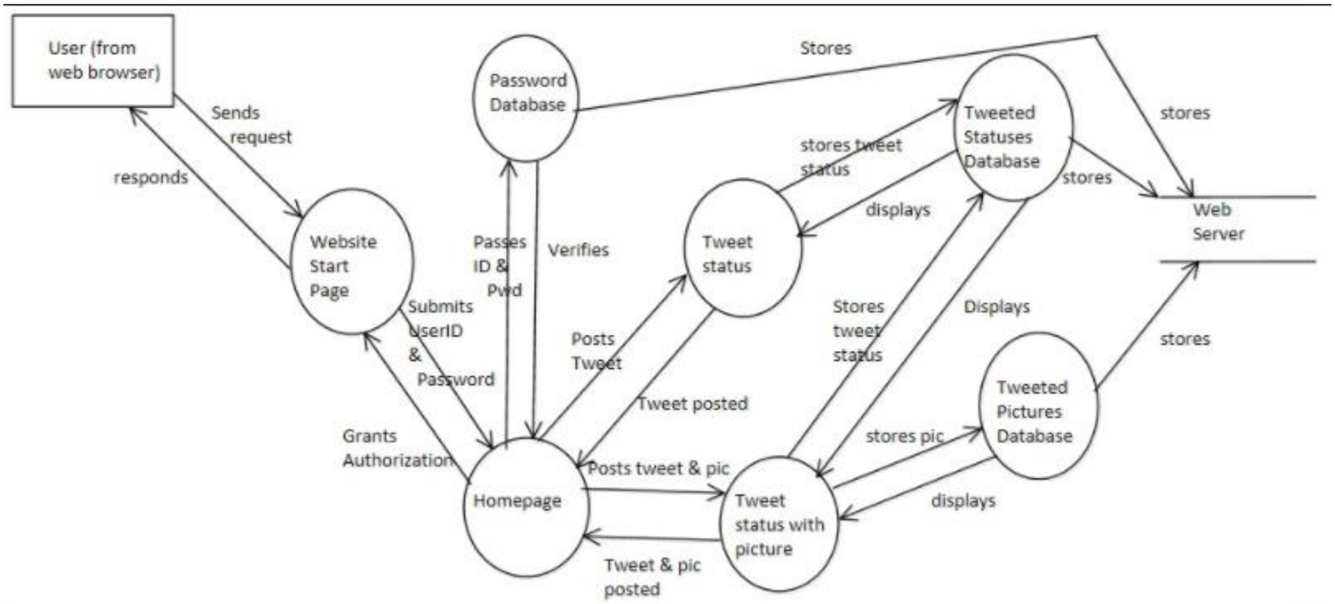


Figure 4.5.3: DFD Level 2

This diagram shows how the system will take the decision. It will first take the input. First stage user will login into the website after that he will tweet on website and when the sentiment analysis of the users tweet is done then result will provide to the website in three category that is positive,negative and stressed.

4.6. Use Case Diagram:

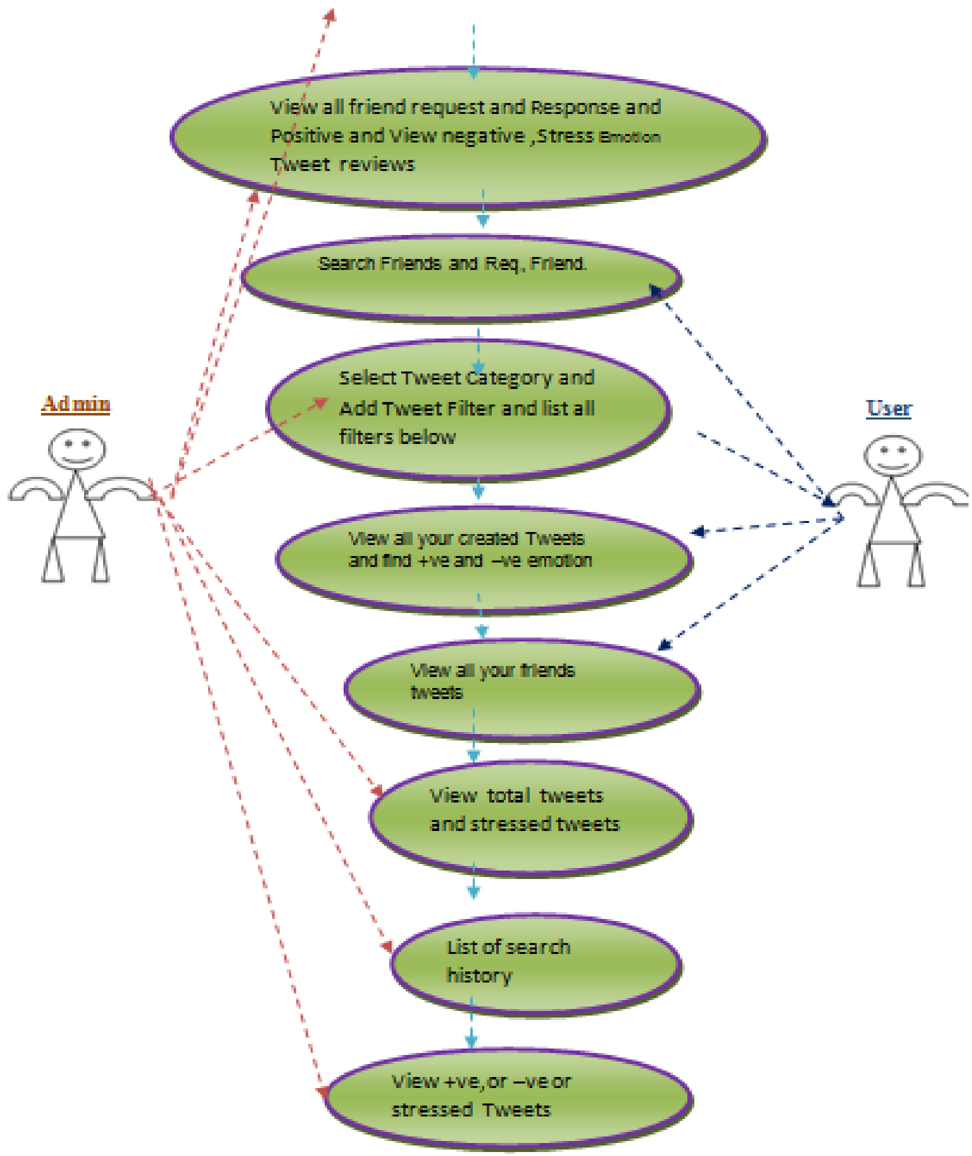


Figure 4.6.1: Use case diagram

This is the use case diagram of detecting stress based on social interactions in social networks.There are two users in our system first is OSN user and other one is Admin user.User can search for friend ,request friend ,create tweet by tweet name ,And find the positive,negative and stressed tweets.Admin user can view all the users requests and add the tweet category like positive,negative ,stressed. Admin user can also view stress emotions.

CHAPTER 5: IMPLEMENTATION

5.1 Implementation steps

- Define the goal. Here, detecting the users stress based on tweets
- Open the website
- Register into the website
- Login into the system
- View all the tweets
- Search friend or request friend
- View all the friends
- Create a tweet by tweet name, tweet description, tweet image and date
- View all the tweets and find all the positive,negative and stress emotions of your tweets.
- Admin can view all the users request and response
- Add the tweet category positive,negative and stressed
- Listing down all the users tweets as well as the emotions of user based on his tweets using sentiment analysis
- Graph of sentiment analysis of particular user is shown as Positive , Negative or Stressed.

CHAPTER 6 : TESTING

6.1 Testing

Software testing is an investigation conducted to provide stakeholders with information about the quality of products or services under test. Software testing can also provide an objective, independent view of the software to allow the business to appreciate and understand the risks of software implementation. Test techniques include the process of executing a program or application with the intent of finding software bugs, errors or other defects.

Types Of Testing

1. Unit Testing:

Unit testing is usually conducted as part of a combined code and unit test phase of the software life cycle, although it is not uncommon for coding and unit testing to be conducted as two distinct phases.

Test strategy and approach

Field testing will be performed manually and functional tests will be written in detail.

Test objectives

- All field entries must work properly.
- Pages must be activated from the identified link.
- The entry screen, messages and responses must not be delayed.

Features to be tested

- Verify that the entries are of the correct format
- No duplicate entries should be allowed
- All links should take the user to the correct page.

2 . Integration Testing

Software integration testing is the incremental integration testing of two or more integrated software components on a single platform to produce failures caused by interface defects.

The task of the integration test is to check that components or software applications, e.g. components in a software system or – one step up – software applications at the company level – interact without error.

Test Results: All the test cases mentioned above passed successfully. No defects encountered.

3. Acceptance Testing

User Acceptance Testing is a critical phase of any project and requires significant participation by the end user. It also ensures that the system meets the functional requirements.

Test Results: All the test cases mentioned above passed successfully. No defects encountered.

6.2.1 Test Case ID: 01

Test Case Description: To check User's authentication using valid username and password.
Modules to be tested: Login.
Expected Result:
If fields are left empty a message "Please fill out this field" is displayed.
If incorrect value is submitted a message "Incorrect input" is shown.
Actual Result: PASS

6.2.2 Test Case ID: 02

Test Case Description: To check whether admin is able to manage the users.

Modules to be tested: Admin module.

Expected Result:

- The admin is able to add and delete the users successfully.
- The admin is able to add and delete the user category such as positive, negative and stressed.
- View all the users tweets
- View user’s emotions.

Actual Result: PASS

6.2.3 Test Case ID: 03

Test Case Description: To check whether the user is able to tweet through the website.

Modules to be tested: User Tweets.

Expected Result:

User is able to post their tweets through the website.

Actual Result: PASS

6.2.4 Test Case ID: 04

Test Case Description: To check whether the user is able to detect stress.

Modules to be tested: Stress detection.

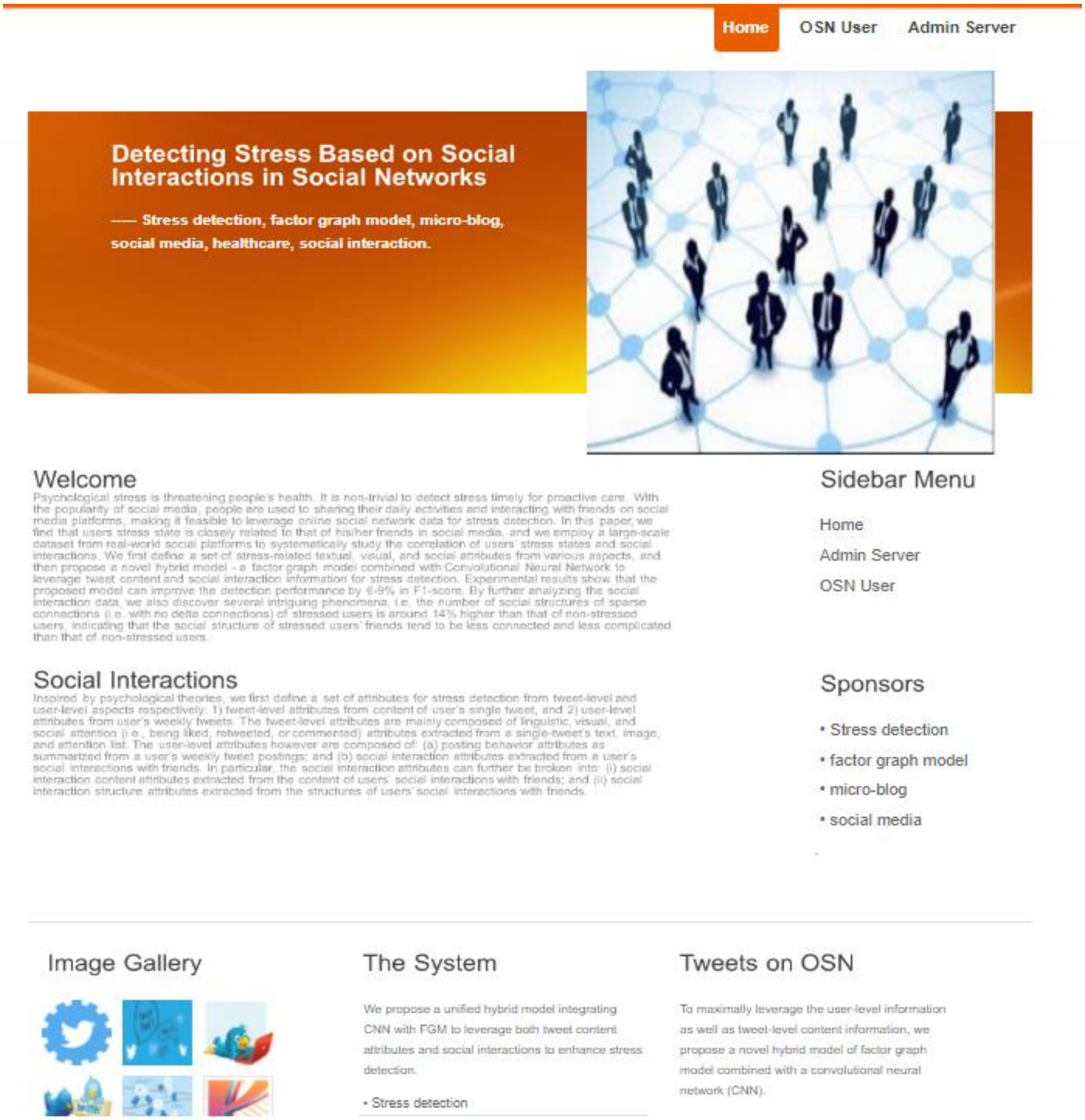
Expected Result:

The user is able to detect stress in three categories such as positive, negative and stressed.

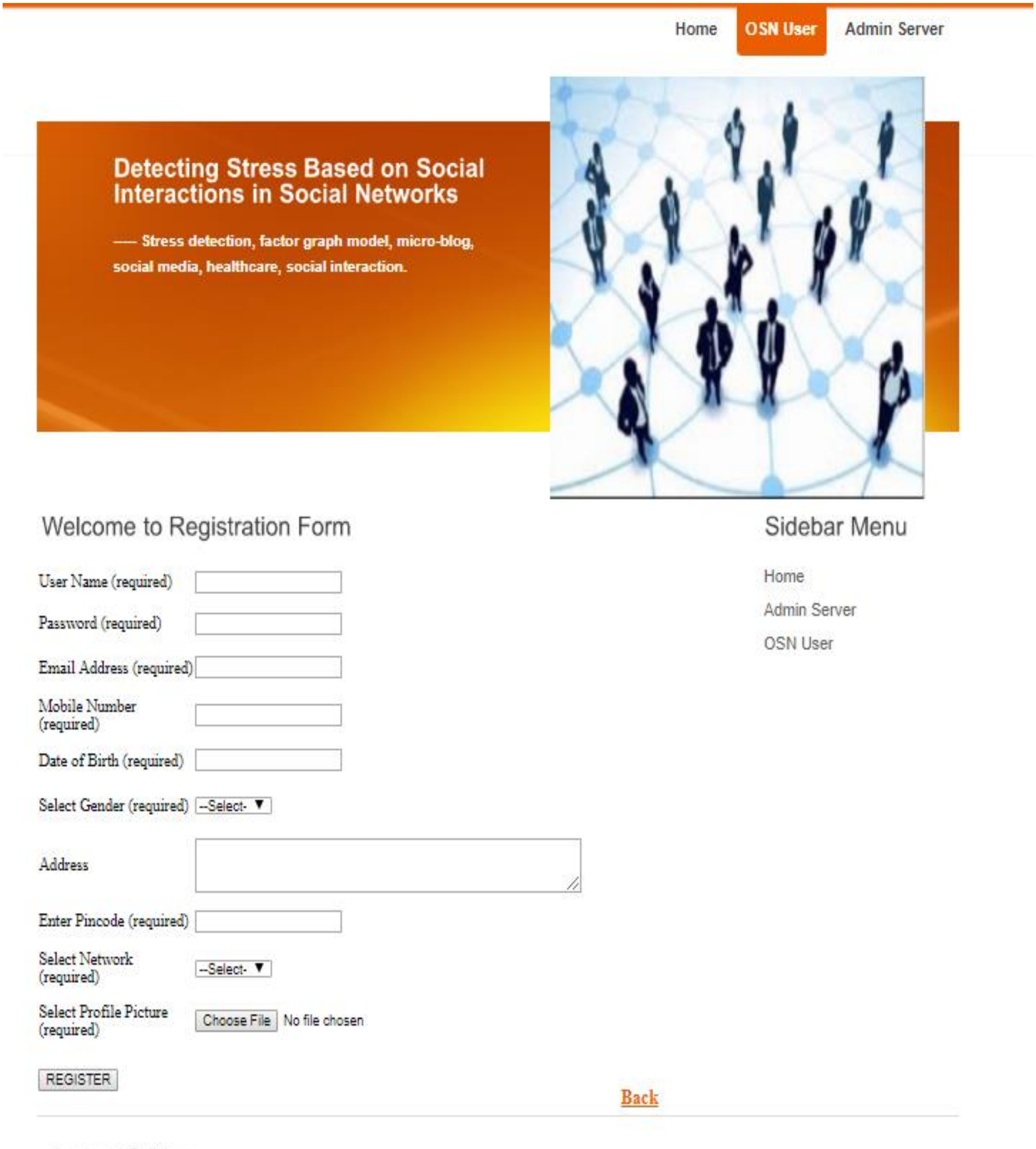
Actual Result: PASS

CHAPTER 7 :RESULT ANALYSIS

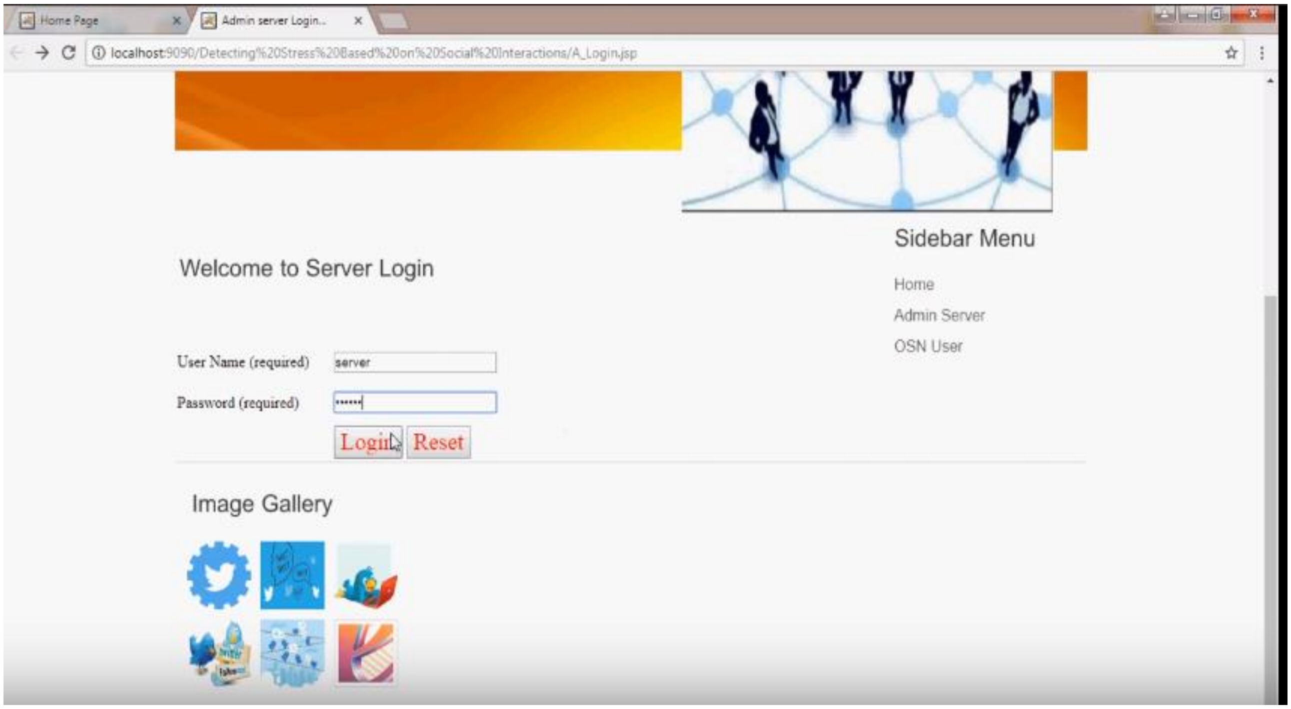
7.1 Output Printouts:



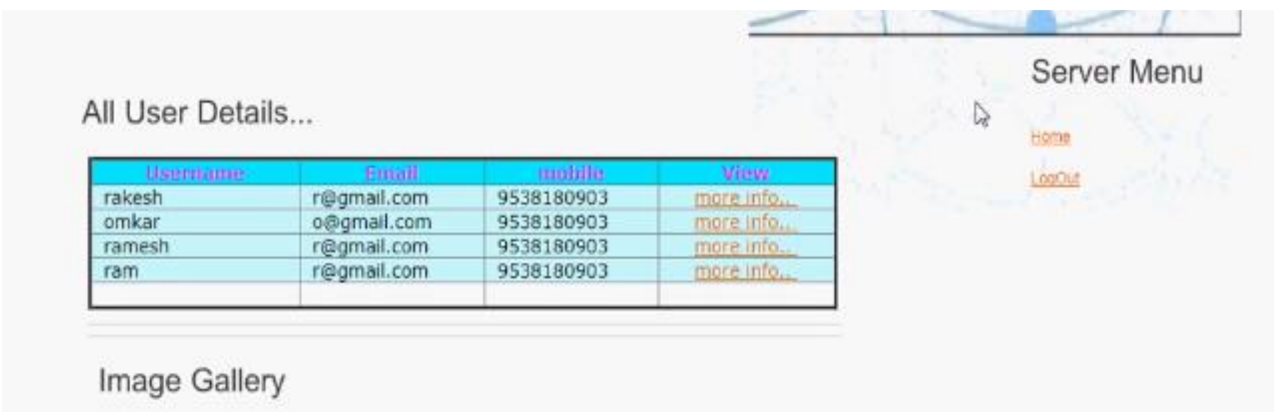
7.1.1 Home screen



7.1.2 : User Register



7.1.3 : Admin login



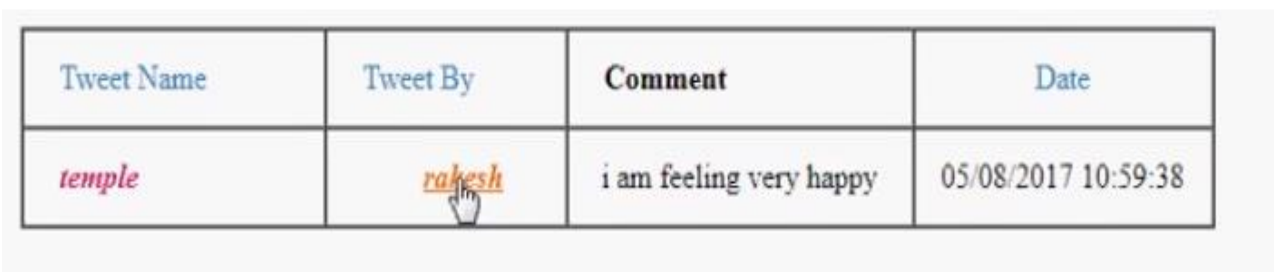
7.1.4 : All User details

7.1.5 : User detail

7.1.6 : Add User category

Image	Tweet Name	Description	Date	Tweets	Re-Tweets
	temple	i am feeling very happy	05/08/2017 10:59:38	View Tweet Details	View Re-Tweet Details
	mountain	i am feeling very sad here	05/08/2017 11:00:11	View Tweet Details	View Re-Tweet Details
	city	i went city and i got stressed	05/08/2017 11:00:55	View Tweet Details	View Re-Tweet Details
	company	to ady i went for company i am feeling so bad	05/08/2017 15:20:00	View Tweet Details	View Re-Tweet Details
	Metro	i am very happy with metro ride	05/08/2017 17:43:01	View Tweet Details	View Re-Tweet Details
	bus	i traveled by bus i am very sad about it	05/08/2017 17:48:52	View Tweet Details	View Re-Tweet Details
	auto	i travelled by auto i am very stressed	05/08/2017 17:50:21	View Tweet Details	View Re-Tweet Details

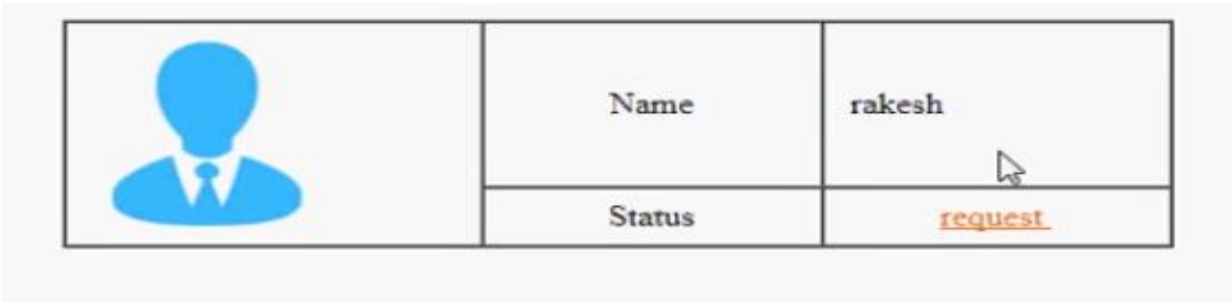
7.1.7 : User Tweets



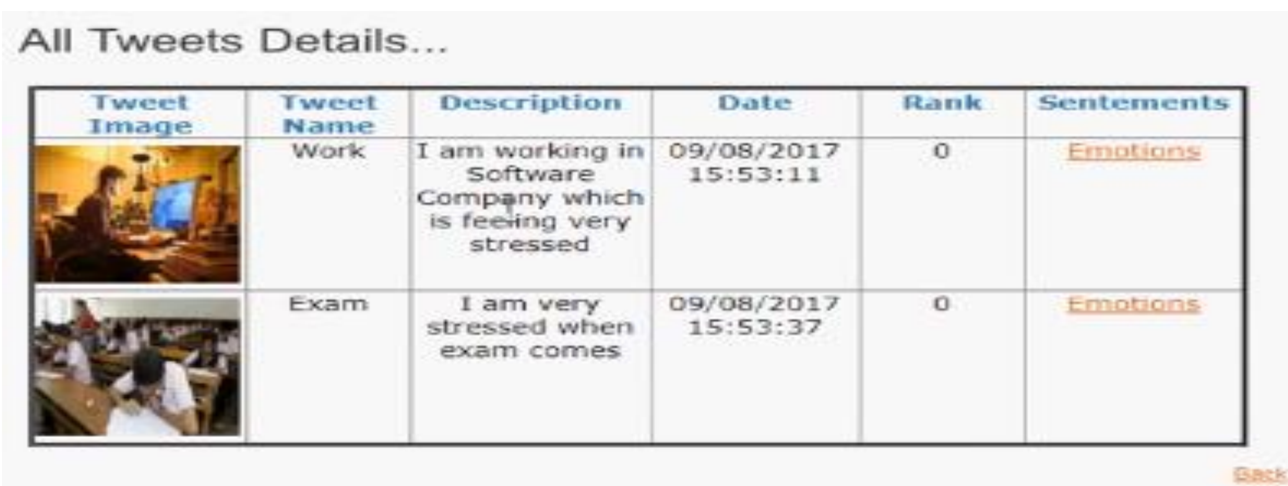
Tweet Name	Tweet By	Comment	Date
temple	rakesh	i am feeling very happy	05/08/2017 10:59:38

7.1.8 : Particular User Tweet

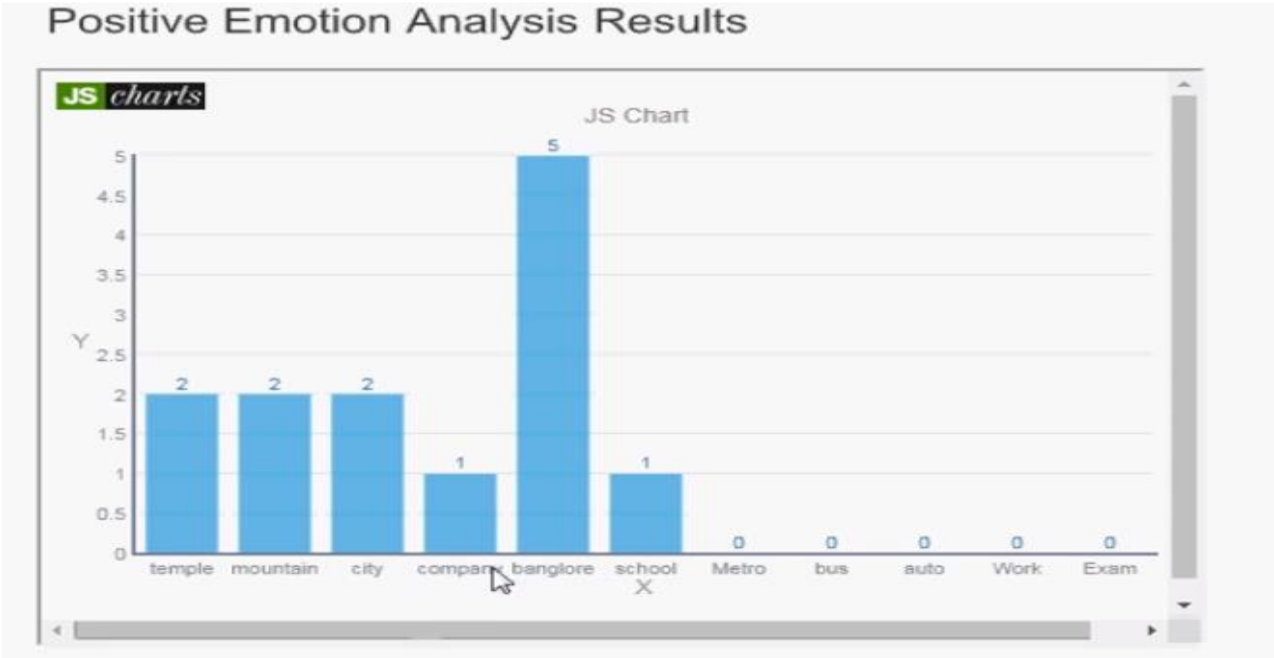
7.1.9 : Search Friend



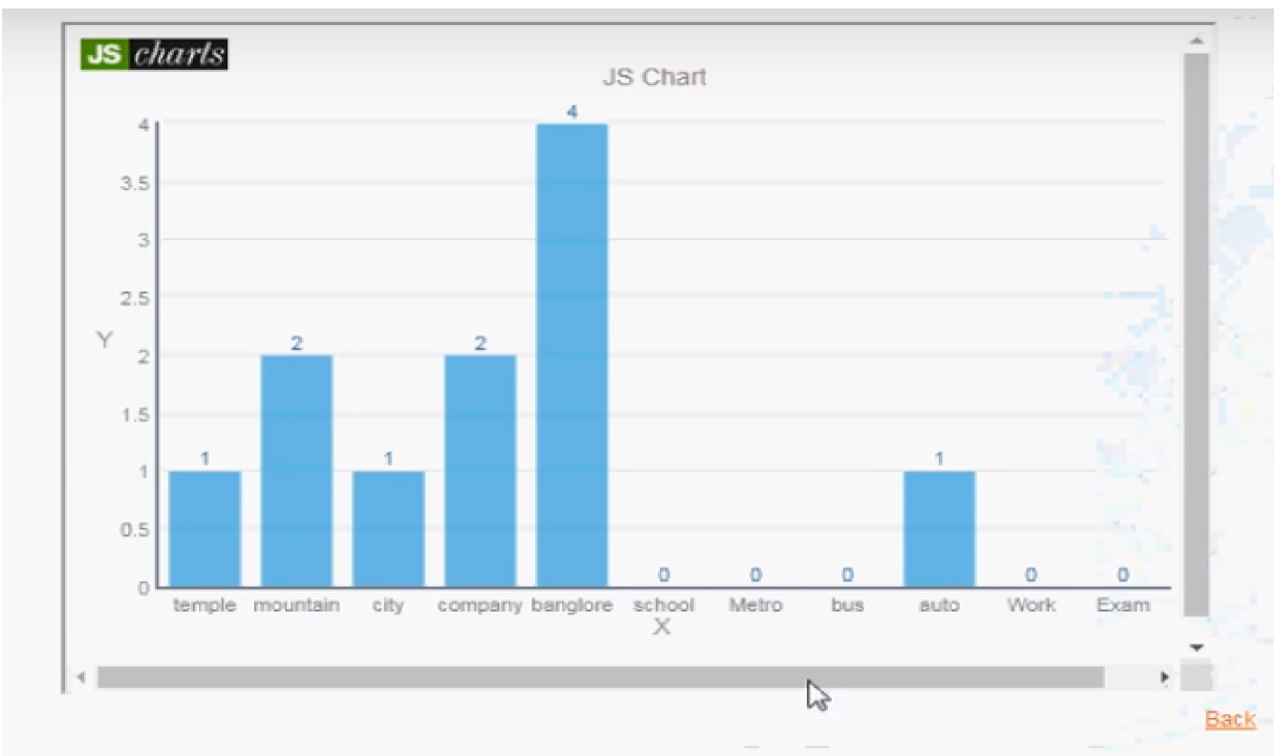
7.1.10 : User Tweets from the website



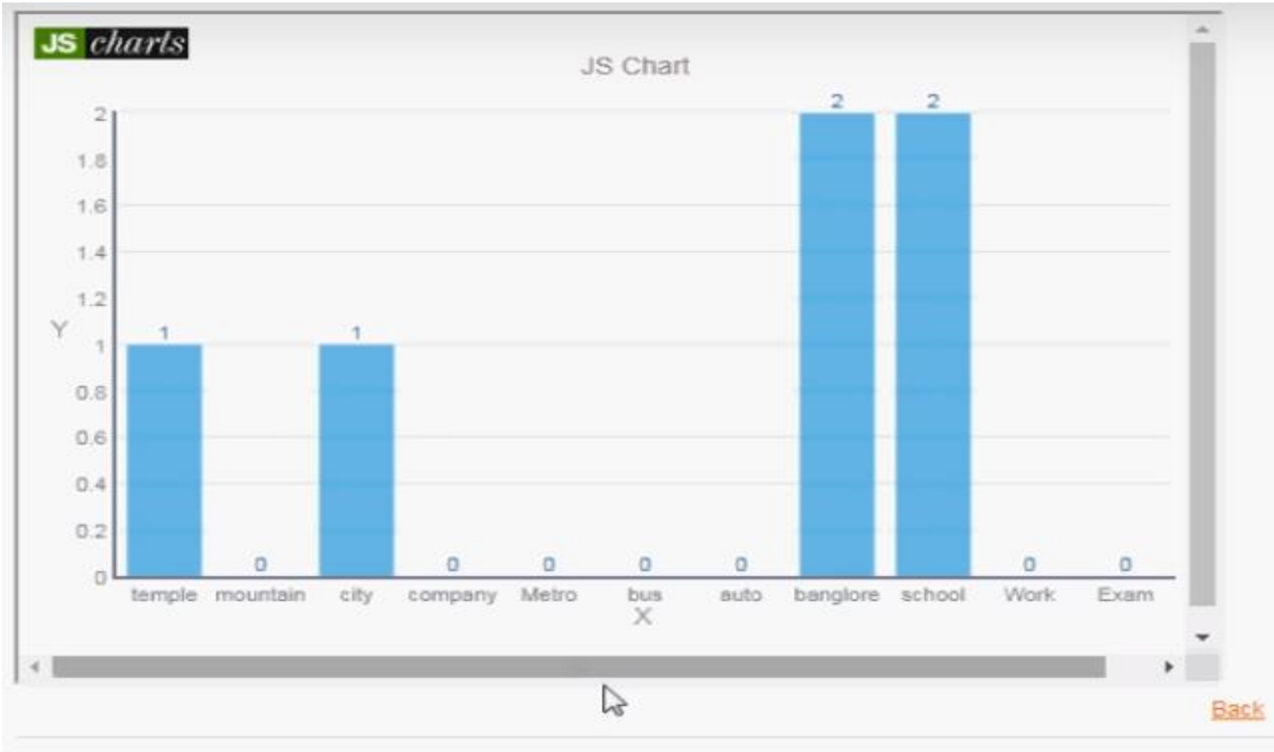
7.1.11 :User Tweets



7.1.12 : Positive analysis report



7.1.13 Negative analysis report



7.1.14 Stress analysis Report

7.4 Observations & Analysis

- User can find out the stress through website
- User Can tweet using this website
- User will search for a friend and request to the particular friend
- Admin can view all the user tweets
- They also able to add the tweet category.
- Admin can accept and reject the request
- Based on the user tweets User can see the stress level
- The graph will be shown to the user in three categories Positive, Negative and Stressed .

CHAPTER 8: CONCLUSION

8.1 Limitations

- There is no option to tweet based on user emotions.
- Can analyze only single user tweets
- It relies on basic level of skills from users.
- Obtain the result only in three category positive, negative stressed..

8.2 Conclusion

In today's world, where mainly the youth and almost all of the population is suffering from surmounting stress, be it because of peer pressure, work load or other domestic tensions; it is very crucial to have a reality check about how stressed a person really is. It is because of this reason that timely detection and prevention of stress is a dire need. We have come up with this project which assists people in scrutinizing the problem of stress. This project will be very beneficial for those who are not so comfortable in opening up about their problems to others. It will help these people get a reality check and may prompt them to reach out and get medical help, just based on their social interactions. We have utilized both human as well as machine learning and applied the concepts of Sentiment Analysis. The main characteristic of this system is its non-invasiveness and fast-oriented implementation in detecting stress when compared with the previous approaches.

8.3 Future Scope

- This website is helpful for normal users as well as doctors as they can find the stress of the user.And the doctor will provide help to reduce the stress of users.
- In future this website can be upgraded to real time stress detection based on user tweets.
- Also the algorithm used can be upgraded to have efficient and less time complexity and high accuracy.
- The scope of the project is in detecting users' psychological stress states from users' weekly social media data, leveraging tweets' content as well as users' social interactions and also in user-level emotion detection in social networks.
- This system further used by the doctors to give advice to all the users who are stressed.
- It will saves the time of doctors as well as users.

References

[1] Andrey Bogomolov, Bruno Lepri, Michela Ferron, Fabio Pianesi, and Alex Pentland. Daily stress recognition from mobile phone data, weather conditions and individual traits. In *ACM International Conference on Multimedia*, pages 477–486, 2014.

[2] Chris Buckley and Ellen M Voorhees. Retrieval evaluation with incomplete information. In *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 25–32, 2004.

[3] Xiaojun Chang, Yi Yang, Alexander G Hauptmann, Eric P Xing, and Yao-Liang Yu. Semantic concept discovery for large-scale zero-shot event detection. In *Proceedings of International Joint Conference on Artificial Intelligence*, pages 2234–2240, 2015.

[4] Wanxiang Che, Zhenghua Li, and Ting Liu. Ltp: A Chinese language technology platform. In *Proceedings of International Conference on Computational Linguistics*, pages 13–16, 2010.

[5] Chih chung Chang and Chih-Jen Lin. Libsvm: a library for support vector machines. *ACM TRANSACTIONS ON INTELLIGENT SYSTEMS AND TECHNOLOGY*, 2(3):389–396, 2001.

[6] Dan C Ciresan, Ueli Meier, Jonathan Masci, Luca Maria Gambardella, and J ¨ urgen Schmidhuber. Flexible, high performance convolutional neural networks for image classification. In *Proceedings of International Joint Conference on Artificial Intelligence*, pages 1237–1242, 2011.

[7] Sheldon Cohen and Thomas A. W. Stress, social support, and the buffering hypothesis. *Psychological Bulletin*, 98(2):310–357, 1985.

[8] Glen Coppersmith, Craig Harman, and Mark Dredze. Measuring post traumatic stress disorder in twitter. In *Proceedings of the International Conference on Weblogs and Social Media*, pages 579–582,2014.

[9] Rui Fan, Jichang Zhao, Yan Chen, and Ke Xu. Anger is more influential than joy: Sentiment correlation in weibo. *PLoS ONE*, 2014.

[10] Zhanpeng Fang, Xinyu Zhou, Jie Tang, Wei Shao, A.C.M. Fong, Longjun Sun, Ying Ding, Ling Zhou, , and Jarder Luo. Modeling paying behavior in game social networks. In *In Proceedings of the Twenty-Third Conference on Information and Knowledge Management (CIKM'14)*, pages 411–420, 2014.