# Chapter 2

**Software Requirements Specification**

The purpose of this software requirements specification (SRS) document is to specify the user goals and tasks that need to be achieved. It also includes the detailed description of the context and requirements- both functional and non- functional which are vital to the successful completion of this project. Apart from these, the SRS document incorporates in itself the constraints and assumptions made during the course of this project.

**2.1 Overall Description**

This section describes the general factors that affect the product and requirements of the system. It provides more detailed description of the user requirements. The system requirements states what the system should do and not how it should be implemented. It also provides a high overview of the system showing some very important functions the system should perform. Also various dependencies among the various user requirements are discussed.

1. Twitter data: In this project we are getting data from famous social network Twitter through twitter API’s using twitter4j package.
2. Training data: We are using training data from Stanford research training data.
3. Classifier: In order to classify positive, negative and neutral sentiments we are using Naïve Bayes algorithm.
4. Web host server: We are putting our data into a web host server so that everyone on the internet can access this application.

**2.1.1** **Product Perspective**

The product is designed to allow the user to enter a keyword and then he will get the corresponding positive negative and neutral tweet counts in the form of pie chart or graph.

**2.1.2 Product Functions**

The product takes the keyword from user, retrieves data from twitter, analyzes and classifies the data using Naïve Bayes approach, finally displays the classified data in the form of positive, negative and neutral tweets.

**2.1.3** **User Characteristics**

User has to visit the website and enter the keyword or topic for which he wants to analyze the data. Based on the keyword entered tweets will be retrieved, analyzed and presented to the user in the form of a pie chart or graph.

**2.1.4** **Constraints**

The product can not able to analyze the tweets with the language other than English as our domain is confined only English language and our training data is also in English itself. We are not confined to a single topic so accuracy is limited because of different tweet methods and usage of slang language in tweets.

* + 1. **Assumptions and Dependencies**

Bag of Words assumption: We assume that the position of words(features) in the document doesn’t matter

Conditional Independence: We assume that the feature probabilities *P*(*xi*|*cj*) are independent given the class *c,* that is

P(x1,x2,……,xn|c) = P(x1|c)\*P(x2|c)\*..\*P(xn|c)

Another assumption is that the user has a basic knowledge about Twitter and it’s orthography that includes special features such as hash tags(#), Usernames(@).

**2.2 Specific Requirements**

In this section, we are going to describe a set of requirements which are specific to the application.

**2.2.1 Functionality Requirements**

The system provides the functionality that is, it will return a set of results based on the keyword that the user has entered and that these results will be classified according to the Naïve Bayes algorithm.

**2.2.2 Performance Requirements**

The system will perform in an efficient way to classify the data and make sure that there are performance enhancements in the form of retrieving the data from Twitter and analysing them with respect to training data.

**2.2.3 Software Requirement**

* Operating System: Windows 7/8, Linux, Mac
* Software Tools : Java, JDK 7.0, Apache Tomcat Server version 7.0
* Web Browser (Mozilla, IE8+, Chrome)
* IDE : Eclipse IDE for J2EE Developers
* API Libraries : Twitter4j library
  + 1. **Hardware Requirement**
* Processor : 1.6 GHz CPU (Minimum),

2.2GHz or higher CPU (Recommended).

* Memory : 384 MB RAM (Minimum),

1. RAM or more RAM (Recommended).

* Hard disk : 5400 RPM (Minimum),

7200 RPM or higher hard disk (Recommended).

### Display : 1024x768 display (Minimum),

### 1280x1024 display or higher resolution display with 256 Colours.