2016

Real – time sentiment analysis for detection of terrorist activities in Kenya

Ngoge, L. A.
Faculty of Information Technology (FIT)
Strathmore University

Real – Time Sentiment Analysis for Detection of Terrorist Activities in Kenya

Lucas Achuku Ngoge

A research thesis submitted in partial fulfillment of the requirements of the Degree of Master of Science in Information Technology at Strathmore University.

Faculty of Information Technology
Strathmore University,
Nairobi, Kenya.

June 2016

DECLARATION

I declare that this work has not been previously submitted and approved for the award of a degree by this or any other University. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.

Name: Lucas Achuku Ngoge

Signature:

Date: June, 2016

Approval

The thesis of Lucas Achuku Ngoge was reviewed and approved by the following:

Dr. Joseph Orero,
Dean, Faculty of Information Technology
Strathmore University

Prof. Ruth Kiraka
Dean, School of Graduate Studies
Strathmore University

ABSTRACT

Terrorism has become a subject of concern to many people in Kenya today. Majority of people are worried lot because they don't know when they will become victims of terrorists' activities. Corruption, porous border and luck of government in the neighboring Somali, have made Kenya a potential target for terrorists'. The advancement in technology has brought a new era in terrorism where Online Social Networks such as Twitter, Facebook has driven the increase use of the internet by terrorist organizations and their supporters for a wide range of purposes including recruitment, financing, propaganda, incitement to commit acts of terrorism and the gathering and dissemination of information for terrorist activities. Although the Kenya government improved its ability to fight terrorism but the changing pattern of terrorist activities, human errors and delayed crime analyses have given criminals more time to destroy evidence and escape arrest.

The evolution of computerized systems has made tracking of terrorist' activities easier. This has helped the law enforcement officers to speed up the process of solving crimes. In this research data was collected from twitter then followed by sentiment analysis on tweets collected to derive rules for the real-time classifier. Geographic analysis was done to reveal a correlation between the tweets and the terrorist' activities as portrayed by the map.

The main objective of this research is to develop a model that will be used to establish crime patterns associated with terrorist activities using sentiment information deduced from twitter data. To achieve this objective, 346 tweets related to terrorism were collected, cleansed and stored in a database for a period of 7 days. This data was then used as features for training and development of the model which will then be used to carry out real time sentiment analysis on twitter data. The model was tested and it was able to classify text correctly into positive, negative and neutral classes with an accuracy score of 73%.

DEDICATION

I dedicate this work to my wife, my son and my two daughters for giving me moral support and good will during my absence in time of study. To all of you, Thank you.

ACKNOWLEDGEMENTS

I take this opportunity to thank the faculty of Information Technology, Strathmore University for giving me a chance to pursue and successfully complete this course.

My sincere thanks go to my supervisor, Dr. Joseph Orero and Prof. Ismael Ateya and the other members of staff in their various capacities for the untiring support, guidance and concern throughout my course work.

I also thank my classmates for the encouragements and team work we have shared during class sessions. Above all, I thank God for his providential care all through the time of my study.

DEFINITION OF TERMS

Dark Websites are websites used by terrorists, radicals and extremist groups for communication and disseminating their ideologies to the public.

Terrorism is an anxiety-inspiring method of repeated violent action, employed by (semi-) clandestine individuals, group or state actors, for idiosyncratic, criminal or political reasons.

Social media is the interaction platform among people in which they create, share or exchange information and ideas in virtual communities and networks.

Twitter is an online social networking and micro blogging service that enables users to send and read short 140-character text messages.

A blog is a discussion or informational site published on the World Wide Web and consisting of discrete entries (posts) typically displayed in reverse chronological order (the most recent post appears first).

Python is a widely used general-purpose, high-level programming language. Its design focuses on code readability, and its syntax allows programmers to express concepts in fewer lines of code than would be possible in languages such as C. The language provides constructs intended to enable clear programs on both a small and large scale.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE: INTRODUCTION

1.0 Background of the study

Kenya is considered to be a beacon of stability and peace in the horn of Africa, greater Eastern Africa region and in Africa, Kenya plays an important role. It serves as an economic and business hub for both national and international investors (Leah & Abdalla, 2009). However, terrorism has threatened its peace and stability. It has undermined the freedom of association and movement of citizens and created a sense of fear and intimidation which has hampered the spiritual, economic and social development of individuals (Leah & Abdalla, 2009).

Terrorism is a violent action employed by extremist groupings for criminal or political reasons. It aims at threatening a particular person or sometimes many people. Terrorists choose their victims because of their symbolic importance. To achieve this goal, they strive to reach large audience. These groupings rely on online social networks (OSNs) to broadcast and transmit their messages to reach many people (Jytte, 2015). The advancement of technology has driven the increase of use of internet in developing countries. Numerous studies have shown that users from developing regions are spending a significant amount of time in online social networks (OSNs) (Fredrick, 2013).

The benefits of Internet Technology are numerous, ranging from its unique suitability for sharing information and ideas to ease of communication with relative anonymity, quickly and effective across border (UN, 2012). However, the same technologies that facilitate communication can be exploited for purposes of terrorism (UN, 2012). For example, Al-Shabaab, a terrorist group based in Somalia, used Twitter during a Westgate terrorist attack in Nairobi to disseminate information and claiming responsibility of the attack (Fredrick, 2013). However, during the Westgate attack Twitter was used to mobilize the country for blood donation, money donation as well as keeping the peace (Fredrick, 2013).

The means by which the internet is often utilized to promote and support acts of terrorism are propaganda (recruitment, radicalization, and incitement to terrorism), financing, training and planning acts of terrorism (UN, 2012). First, Propaganda; it takes the forms of multimedia communications providing ideologies or practical instructions or explanations or justifications for promotion of terrorist activities. Second, financing; terrorists use the Internet to raise and

collect funds and resources (UN, 2012). Third, training; internet provides a platform for detailed instructions, often easily accessible multimedia format and multiple languages, on topics such as how to join terrorist organizations, how to construct explosives, firearms or other weapons or hazardous materials and how to plan and execute terrorist attacks (UN, 2012). Finally, planning; it is an act of terrorism typically involves remote communication among several parties. Internet technology facilitates the preparation of acts of terrorism, through communications within and between organizations promoting violent extremism, as well as across borders (UN, 2012).

Twitter has recently emerged as terrorists' favorite Internet service, even more popular than self-designed websites or Facebook, to disseminate propaganda and enable internal communication. For example, without Twitter, the explosive growth of ISIS over the last few years into the most-feared terrorist group in the world would not have been possible (Weimann, 2014). In addition, the militant group Al-Shabaab, during its September 2013 attack on Westgate Mall in Nairobi, Kenya, gave a live commentary on its actions on Twitter.

Terrorists have good reasons to use twitter. First, it is by far the most popular with their intended audience, which allows terrorist organizations to be part of the mainstream. Second, twitter APIs are user-friendly, reliable, and free. Finally, Twitter APIs allows terrorists to reach out to their target audiences and virtually "knock on their doors" in contrast to older models of websites in which terrorists had to wait for visitors to come to them. (Weimann, 2014). While terrorists have developed many ways to use the Internet in furtherance of illicit purposes, their use of the Internet also provides opportunities for the gathering of intelligence and other activities to prevent and counter acts of terrorism, as well as for the gathering of evidence for the prosecution of such acts (UN, 2012). For example, the Kenya government improved its ability to fight terrorism by increasing its capabilities to identify, arrest and detain suspects through an Anti-Terrorism Police Unit (ATPU), Yet al-Shabaab's advances in Somalia have challenged Kenya's ability to prevent terrorist attacks at home (Samuel, 2013).

The Kenya government initiated counter-terrorism measures in order to combat the imminent threat in her bid to preempt, foil and counter any attacks on the citizens (Samuel, 2013). However, these techniques have not been effective in counter terrorists' activities largely because first, the Anti-Terrorism Police Unit (ATPU) is too small and not effective to counter a growing extremist minority in Kenya (Samuel, 2013). Finally, these methods are untimely and have a lot of organizational and bureaucratic limitations which prevent the intelligence and

national security organizations from sharing vital information with other agencies (Samuel, 2013).

Increased Internet use for terrorist purposes provides a corresponding increase in the availability of electronic data which may be compiled and analyzed for counter-terrorism purposes. Law enforcement, intelligence and other authorities are developing increasingly sophisticated tools to proactively prevent, detect and deter terrorist activity involving use of the Internet (UN, 2012).

Therefore, sentiment analysis can be used to help with identification, detection and prevention of terrorists' activities in an actionable and timely manner. Sentiment analysis is the means of applying natural language processing methods and determining subjective information in source text (Vishal & Sonawane, 2016). It is one of most important source for decision making. The sentiment analysis techniques can be used to extract, identify, evaluate and classify online sentiments (Vishal & Sonawane, 2016). It involves identification of the overall mood, feeling and speculation of a text (Pravesh & Mohd, 2014).

1.1 Problem statement

The current terrorist' activities detection methods are untimely and have a lot of organizational and bureaucratic limitations which prevent the intelligence and national security organizations from sharing vital information with other agencies thus making it ineffective to counter terrorists' activities. This situation is coupled by the collusion between the security personnel and members of the clandestine organizations resulting to their victory thus jeopardizing government's effort to restore security (NCRC, 2012).

Human investigators who have a lot of experience in intelligence gathering, always analyze crime patterns precisely however as complexity of crime increases, human errors and increase in analysis time, have given criminal gangs more time to destroy evidence and escape police arrest (Hsuchun et al, 2004).

Hence, there is need for adoption of real-time sentiment analyzer to help with identification and detection of terrorists' activities and provide summaries of crime data which are intended to find their applicability in supporting law enforcement agencies in mitigating terrorist threats.

1.2 Research objectives

i. To investigate the current terrorism detection methods in use.

ii. To identify terrorist activities using the current terrorism detection methods

iii. To review the existing data mining techniques available for use in crime detection.

iv. To develop a prototype that can be used for detection of terrorist activities in Kenya.

v. To test the prototype.

1.3 Research questions

i. What are terrorism detection measures currently employed?

ii. What are the gaps associated with the current counter terrorism techniques used for crime detection?

iii. How can machine learning techniques be used in analyzing terrorist activities?

iv. What is the most efficient machine learning techniques used for crime detection?

v. Does the prototype works as expected?

1.4 Justification

The objective of this research is to come up with a prototype that will help the law enforcement officers in mitigating terrorist threats by classifying and mapping sentiments expressed in social media on a map. This will help law enforcement officers in solving crimes faster going by their distribution on a map. It will also give the law enforcement officers investigative leads and information which will help them in disrupting, exposing and uncovering terrorists' networks and their structure effectively which would otherwise take hours to uncover manually.

1.5 Limitations

The words used on social media especially Facebook, Twitter etc., do not constitute formal language. They involve acronyms, emoticons and slang. This at times makes it difficult to classify sentiments. Secondly identifying the hardware, from which the tweets are sent, is a challenge. Finally knowing the threshold of words that can be used by law enforcers to prosecute a criminal is very difficult. Due to the limitation of time for this study only small size of twitter data is used in the experiment.

CHAPTER TWO: LITERATURE REVIEW

2.0 Introduction

As outlined in the introduction, this research explores the use of sentiment analysis in developing a prototype for detection of terrorist activities in Kenya. This section presents a literature review of crime pattern analysis and machine learning techniques available for use in crime data mining. Online social media has improved the way many people do businesses as well as interact with each other. Twitter has become a platform of choice by many people because of its ability to broadcast small pieces of information to a large number of people. It is therefore an effective form of mass communication. The ease of communication has allowed the public to freely exchange anything they wish through the platform (James, 2012).

The social media has both benefit and harm in a number of ways. For example, communities and some law enforcement agencies have been creating and encouraging the use of social networks to create a 'virtual neighborhood watch' that can consist of crime alerts from law enforcement organizations in order to counter criminal activities. However, terrorist organizations have taken advantage of its availability to disseminate their ideologies to the public and eventually carry out terrorist attacks (Jytte, 2015).

2.1 Crime pattern analysis

Crime analysis is a law enforcement function that involves the systematic analysis of identifying and analyzing both pattern and trends in crime and disorder (Bolla, Raja & Ashok, 2014). Criminal pattern analysis is very crucial in combating crime. Computer systems have to be engaged in order to gather and interpret intelligence so as to control the criminal environment as well as influence effective decision making (Kester, Quist-aphetsi & Mieee, 2013). Figure 2.1 shows stages in criminal analysis. Analysis of a crime pattern typically focuses on who, what, when, where, and how factors that are common across a significant number of incidents are identified. Identification of these commonalities is often the key to finding solutions to criminal activities (IACA, 2008).

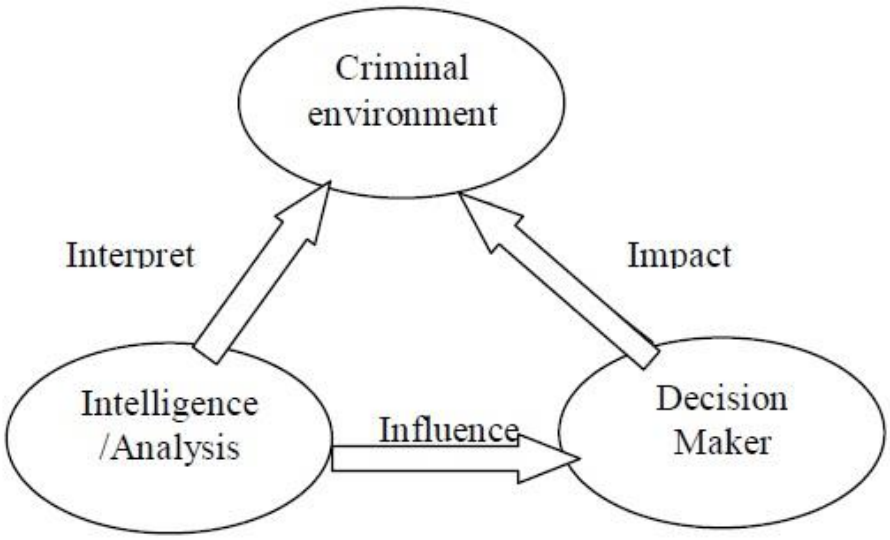

Figure 2.1: Crime analysis process

The difficulty in defining crime analysis is tied to how crime analysis positions are structured and used within individual law enforcement agencies (IACA, 2008). Many analysts in small agencies provide every type of law enforcement analysis for their organizations, but it is rare to find these analysts titled "law enforcement analyst". Instead, the agencies call them "crime analysts" or "intelligence analysts" (IACA, 2008).

The crime analysis process begins with collecting and managing data thereafter the actual analysis of a crime pattern begins with its initial identification of criminal activities and their patterns. To accomplish crime analysis, the analysts must ensure that he has real-time access to incident reports and make information scanning a daily process (Kester, Quist-aphetsi & Mieee, 2013). The analyst can determine the criminal patterns based on one of three factors: first, the modus operandi commonalities, found through a careful review of incident reports and their narratives; second, Exceptional Volume, found through some brand of threshold analysis, either deliberate or unconscious; and lastly, Geographic Proximity, found through crime mapping.

2.1.1 Review of incident reports

Report reviews should be done on a daily basis. The lack of time or manpower sometimes makes it impossible to accomplish. This process is the least scientific of all the means to identify patterns, but it is probably the most effective. However, in Kenya, the law enforcement officers

get it difficult to review incident reports; this is mainly due to the fact that many crimes go unreported. These incidences are also handled through alternative dispute resolution mechanisms in the villages and communities (ICPC, 2008). The validity of police reports has been questionable with media reports increasing unaware of crime incidences. However, police records indicate a decline in these incidences (Leah & Abdalla, 2009). The terrorist and extremist groups have taken advantage of the gaps existing in security to recruit members of the public through the social media preferable Twitter (Jytte, 2015).

During crime analysis process, the analysts may obtain the results of his query, and then he begins to compare the various factors of each past crime with those of the present one, looking for commonalities that would indicate the existence of a crime pattern (Kester, Quist-aphetsi & Mieee, 2013). Again, there is no scientific formula for the analyst to apply in this process; he must use his own knowledge and experience to decide at which point one or more commonalities indicates the likelihood of a pattern. Such a determination depends, again, on the nature of crime in the analyst's individual jurisdiction.

It is noted that, some of the efforts have been made to automate the reports review process, automatically scan for similarities among incidents, and identify potential patterns without requiring the analyst to read each report. While sophisticated data mining and neural network technology are promising, it has not achieved a lot in identification of criminal activities enough to make the analyst's daily review redundant regular basis.

2.1.2 Threshold Analysis

Threshold analysis describes the process by which the analyst identifies potential patterns through exceptional volume (Kester, Quist-aphetsi & Mieee, 2013). The theory behind threshold analysis is that when crime in a particular geographic area reaches a level that is significantly higher than usual, some type of crime pattern is probably to blame. The analyst can use a statistical method to determine when crime has reached a level that is "significantly higher than usual" in other words, when crime crosses the threshold from average volume to exceptional volume. However, absent any statistical methods, an experienced analyst will employ an almost unconscious method of threshold analysis.

For example, in Isiolo, the Kenya government introduced mobile phones to enhance the ability of community members to communicate with law enforcement officers in the area (ICPC, 2008) while in Kibera, the government established the so-called "Drop Boxes" to encourage people to report crimes and send tip-offs to the police (ICPC, 2008). However, these strategies failed to achieve much as crime and terrorists' activities are still on the rise (ICPC, 2008).

It should be noted that many patterns involve only a few incidents and may therefore not set off any threshold alarms. These small patterns can easily be buried within high average numbers for a particular area. Because of this, threshold analysis should not be the only means by which an analyst seeks crime patterns. The benefit of threshold analysis is that it may catch patterns overlooked in the details of day-to-day report review, and it may benefit understaffed crime analysis units that simply do not have time to effectively use the report review method.

2.1.3 Crime Mapping

The crime mapping method of pattern identification involves creating pin maps or thematic maps of various crimes, and seeking geographic hot spots or clusters. This method is valid only for crime patterns that exhibit geographic clustering (Kester, Quist-aphetsi & Mieee, 2013). Many crime patterns involve crimes that are not geographically close, so, like threshold analysis, the crime mapping method should not be the only means by which an analyst seeks patterns. However, like threshold analysis, mapping crime may help catch patterns lost in the detail of daily report review, and it may be beneficial for understaffed agencies that cannot effectively use daily report review.

2.2 Sentiment analysis

Sentiment analysis is the study that analyses people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions towards entities such as products, services, organizations, and their attributes. Sentiment analysis requires training document of textual content or a data corpus, which serves as a preparation document for classification. The basic machine learning techniques available for text classifications are naive bayes (NB), support vector machines (SVM), maximum entropy amongst others (Pang & Lee, 2008). The law enforcement officers can use these machine learning techniques to predict terrorist activities by analyzing factors such as time, location, address, physical characteristics etc. from a body of a text.

2.3 Sources of opinions

The sources of opinions contributed by people are outlined as follows:

2.3.1 Web discourse or blogs

A blog is an informational site published on the World Wide Web which consists of posts displayed in reverse chronological order (the most recent post appears first). Until 2009 blogs were usually the work of a single individual, occasionally of a small group, and often covered a single subject. More recently multi-author blogs (MABs) have been developed, with posts written by large numbers of authors and professionally edited. The rise of Twitter and other micro blogging systems has helped to integrate MABs and single-author blogs into societal new streams (Govindarajan &Romina, 2013).

2.3.2 Computer-supported collaboration

This type social media focuses on technology that affects groups, organizations, and communities e.g. voice mail and text chat. It grew from cooperative work study that supports people's work activities and working relationships. Network technology increasingly supports a wide range of recreational and social activities, this has expanded consumer markets, enabling more people to connect online to create what researchers call a computer supported cooperative work, which includes all contexts in which technology is used to mediate human activities such as communication, coordination, cooperation, competition, entertainment, games, art, and music (Govindarajan & Romina, 2013).

2.3.3 News Articles

The websites like www.thesun.co.uk, www.cnn.com and www.thehindu.com have news articles that allow users to post their comments. Sites under this category allow a user to interact by voting for articles and commenting on them. It includes all the digital newspapers (Govindarajan & Romina, 2013).

2.3.4 Reviews

There are many user generated reviews available on the internet that aids a customer in buying a product. E-commerce sites such as www.amazon.in, www.flipkart.com and www.reviewcentre has millions of customer reviews for products (Govindarajan & Romina, 2013).

2.3.5 Social Media

Social media differs from traditional and conventional media in many aspects, such as in interactivity, reach, frequency, usability, immediacy, and permanence. It enables anyone to publish, access information, share, co-create, discuss, and modify content (Weimann, 2014). With social media, virtual communities are increasingly popular all over the world. For example, Al-Qaeda, its affiliates and other terrorist organizations have moved their activities to online platform (Weimann, 2014).

2.4 Machine learning methods for sentimental analysis

Sentimental analysis approaches use different machine learning classifiers and feature extractors. In this context, the goal of machine learning is to study the algorithms that are capable in fully automated situations to predict something out of input. There are many ways to do this i.e. the use of Naive Bayes, support vector machines (SVM), maximum entropy etc. There are several applications that have been developed using these algorithms for example; Microsoft Cern with inbuilt random forest for predicting body parts, given what is on the sensor of the camera. Many prototypes and models for sentiment classification treat classifiers and feature extractors as two distinct components (Pang & Lee, 2008).

2.4.1 Naïve bayes classifier

(Pang &Lee, 2008) describes naive Bayes classifier as a supervised machine learning algorithm with a simple probabilistic classifier based on bayes' theorem with strong independence assumptions. The classifier assumes the presence (or absence) of a particular feature of a class is unrelated to the presence (or absence) of the other feature. It can learn the pattern by examining a set of documents that has been categorized. It compares the contents with the list of words to classify the documents to their right category or class (Vishal & Sonawane, 2016). Let d be the tweet and c* be a class that is assigned to d, where

$$C^* = \arg mac_c P_{NB}(c \mid d)$$

$$P_{NB}(c \mid d) = \frac{(P(c))\sum_{i=1}^{m} p(f \mid c)^{n_{i(d)}}}{P(d)}$$

From the above equation, *"f"* is a feature, count of feature "(*fi)"* is denoted with *ni (d)* and is present in *d* which represents a tweet. Here, *m* denotes number of features. Parameters *P(*c) and *P(f|c)* are computed through maximum likelihood estimates, and smoothing is utilized for unseen features. Python NLTK library can be used to train and classify a text using Naïve Bayes (Vishal & Sonawane, 2016).

2.4.2 Maximum entropy

Maximum entropy is a technique for estimating probability distributions from data. In text classification, maximum entropy estimates the conditional distribution of the class label given a document. A document is represented by a set of word count features. The labeled training data is used to estimate the expected value of these word counts on a class-by-class basis (Govindarajan & Romina, 2013).

The principle in maximum entropy is that when nothing is known, the distribution should be as uniform as possible, that is, have maximal entropy. Labeled training data is used to derive a set of constraints for the model that is characterized by the class-specific expectations for the distribution. Using maximum entropy model, prediction of outcome is based on everything that is known and assumes nothing about unknown.

2.4.2.1 Maximum entropy for text classification

In Maximum Entropy Classifier, no assumptions are taken regarding the relationship in between the features extracted from dataset. This classifier always tries to maximize the entropy of the system by estimating the conditional distribution of the class label. Maximum entropy even handles overlap feature and is same as logistic regression method which finds the distribution over classes (Vishal & Sonawane, 2016). Maximum entropy makes no independence assumptions for its features, like Naive Bayes. The model is represented by the following:

$$P_{ME}(c \mid d, \lambda) = \frac{\exp[\sum_i \lambda_i f_i(c, d)]}{\sum_c \exp[\sum_i \lambda_i f_i(c, d)]}$$

Where c is the class, d is the tweet and λi the weight vector. The weight vectors decide the importance of a feature in classification (Vishal & Sonawane, 2016).

2.4.3 Support Vector Machine (SVM)

SVM is a supervised learning model which analyzes the data and identifies the pattern for classification. The concept of SVM algorithm is based on decision plane that defines decision boundaries. A decision plane separates group of instances having different class memberships. For example, consider an instance which belongs to either class Circle or Diamond. There is a separating line (figure 2.2) which defines a boundary. At the right side of boundary all instances are Circle and at the left side all instances are Diamond (Pravesh & Mohd, 2014).

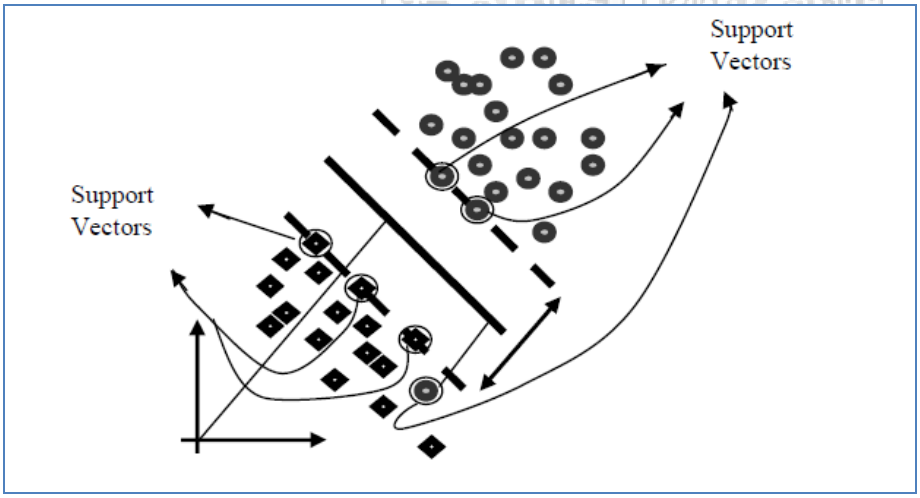

Figure 2.2: Principle of SVM

In text classification sometimes data are linearly divisible, for very high dimensional problems and for multi-dimensional problems data are also separable linearly. Generally, (in maximum cases) the opinion mining solution is one that classifies most of the data and ignores outliers and noisy data. If a training set data say D cannot be separated clearly then the solution is to have fat decision classifiers and make some mistake (Pravesh & Mohd, 2014). The SVM can be used to extract terrorist entities from a collection of untagged news documents in the terrorist domain. This method segments each document into sentences, parses the latter into parse trees and delivers features for the entities within the documents.

2.5 Lexicon-Based Approaches

Lexicon based method uses sentiment dictionary with opinion words and match them with the data to determine polarity (Vishal & Sonawane, 2016). They assign sentiment scores to the opinion words describing how positive, negative and objective the words contained in the dictionary. Lexicon-based approaches mainly rely on a sentiment lexicon, i.e., a collection of known and precompiled sentiment terms, phrases and even idioms, developed for traditional genres of communication, such as the opinion finder lexicon (Vishal & Sonawane, 2016).

2.5.1. Dictionary-based

It is based on the usage of terms (seeds) that are usually collected and annotated manually. This set grows by searching the synonyms and antonyms of a dictionary (Vishal & Sonawane, 2016). An example of that dictionary is WordNet, which is used to develop a thesaurus called SentiWordNet. However, Dictionary-based approach can't deal with domain and context specific orientations (Vishal & Sonawane, 2016).

2.5.2 Corpus-Based

The corpus-based approach provides the dictionaries related to a specific domain. These dictionaries are generated from a set of seed opinion terms that grows through the search of related words by means of either statistical or semantic techniques (Vishal & Sonawane,2016).

2.6 Sentiment analysis processes

The sentiment analysis process involves data collection, text preparation, sentiment detection, sentiment classification and presentation of output (Pang & Lee, 2008).

2.6.1 Data collection

The data collection involves the collection of tweets according to some crime related topics for example use of keywords to identify text/tweets. The keywords such as," Gun", "crime", "kill", "Allah", "kaffir", "attack", "blast", "bomb", "claim", "defend", "force", "kill", "protect", "resist", "strike", etc. is used to track tweets from the twitter (Pang & Lee, 2008).

2.6.1.1 Tweet dataset

Tweet Dataset is a brief details of dataset used in experiments. Dataset needs to be cleaned and tokenized and should be in a CSV format with the headers such as id, date, tweet and location. An example of tweet Dataset is shown in the table 2.1 below.

Table 2.1: Tweet dataset format

ID	Date	Tweet	Location	Latitude	Longitude
12112	Tue Mar 08 16:23:23 +0000 2016	US Bombs Al-Shabaab Terror Camps Graduation ;killing over 150 Terrorists .Https://t.coKn7AMJFrNF	Embu	-0.425	37.531
12182	Tue Mar 08 16:23:46 +0000 2016	Terror fight neeeds multifacet Strategy .Https.//t.co/LPwLiMiY6P			
12100	Tue Mar 08 16:23:04 +0000 2016	Video:Elephant rescued from a well in Kenya .Https./t.c/m^TuZKOxiH	Narok	-1.24076	35.7356

2.6.2 Text preparation

Text preparation involves removal of any unwanted text from tweets collected from twitter and convert them into lower case before classification is performed i.e. removal of alphanumeric characters i.e. hash, dashes **(**Pang & Lee, 2008). Table 2.2 shows a sample of unwanted text.

Table 2.2: Sample of unwanted content and action

S/No.	Unwanted Conten	Wanted Content	Action
1.	#word	Word	Replaced
2.	@username	AT_USER	Converted
3.	https://	URL	Converted
4.	Additional White Space ‘ ‘	“ ”	Removed
5.	Retweet	RT	Removed
6.	Uppercase	lowercase	Converted

2.6.3 Sentiment detection

Sentiment detection requires that each sentence is examined for subjectivity using unigrams and bigrams thereafter a geographical analysis is to be conducted that is, collecting and scrutinizing every sample in a set of samples from which samples are drawn **(**Pang & Lee, 2008).

2.6.4 Sentiment classification

It involves running a sentiment classifier on each extracted sentence to determine if it is positive, negative, or neutral. The basic classifiers available for text classification are naives Bayes, support vector machine, maximum entropy amongst others (Pang & Lee, 2008).

2.7 Crime data mining Methods

Crime data mining aims at identifying patterns from structured and unstructured data for example collected twitter data. Some of the common methods for crime data mining being used are; entity extraction, clustering, deviation detection, classification amongst others.

2.7.1 Entity extraction

This method seeks to identify a particular pattern from text. It is used to identify people, vehicles, and addresses from police narrative reports. The investigators can use this method to analyze the behaviors of serial offenders. However, this method provides very basic information for crime analysis because its performance relies on the availability of large amount of input data (Hsuchun et al, 2004). An example of an application that uses this method is the advanced terrorist detection system (ATDS). This application tracks down terrorist-generated sites, by analyzing the websites visited by the users. Figure 2.3 and figure 2.4 show how ATDS works.

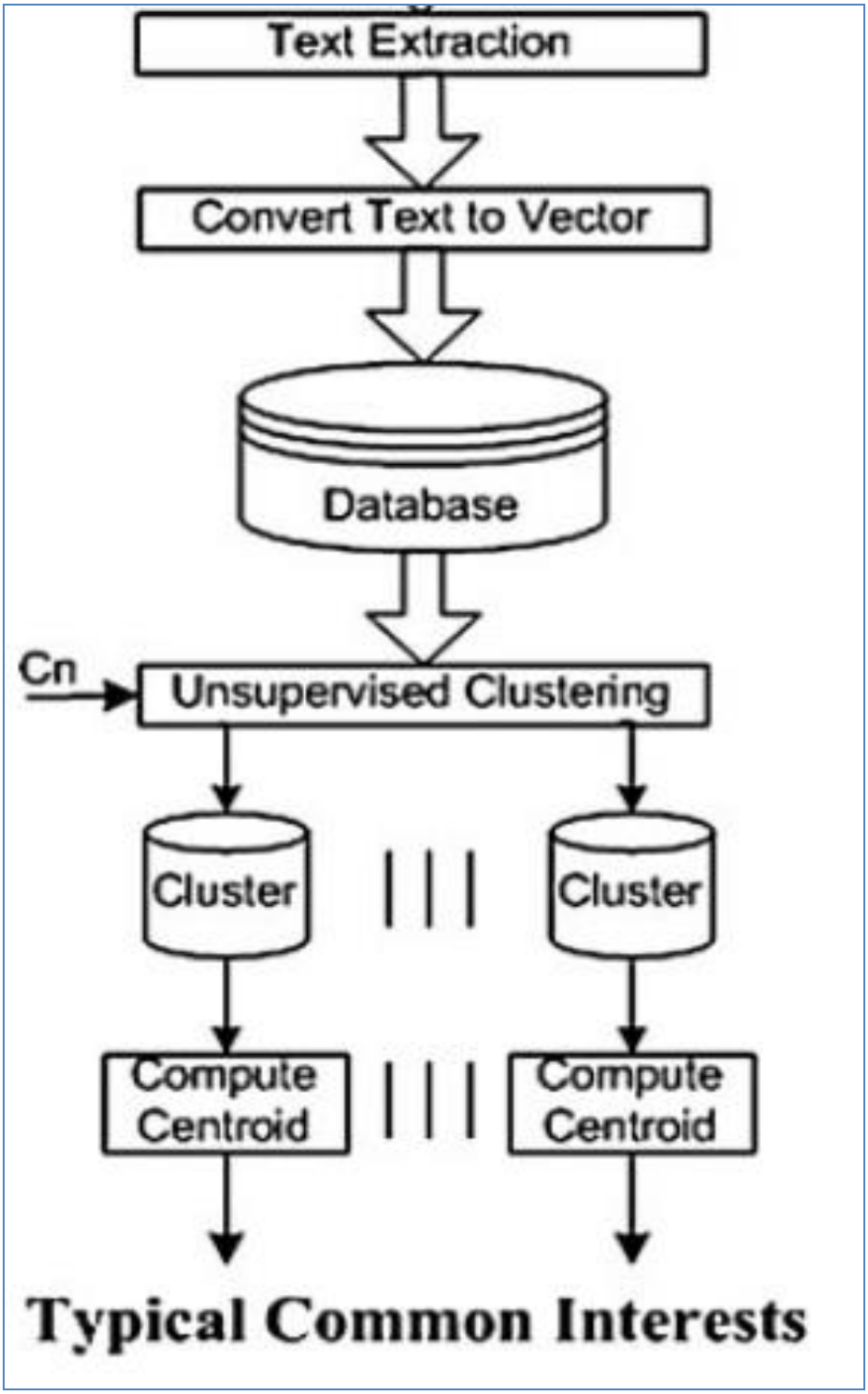

Figure 2.3: Learning mode

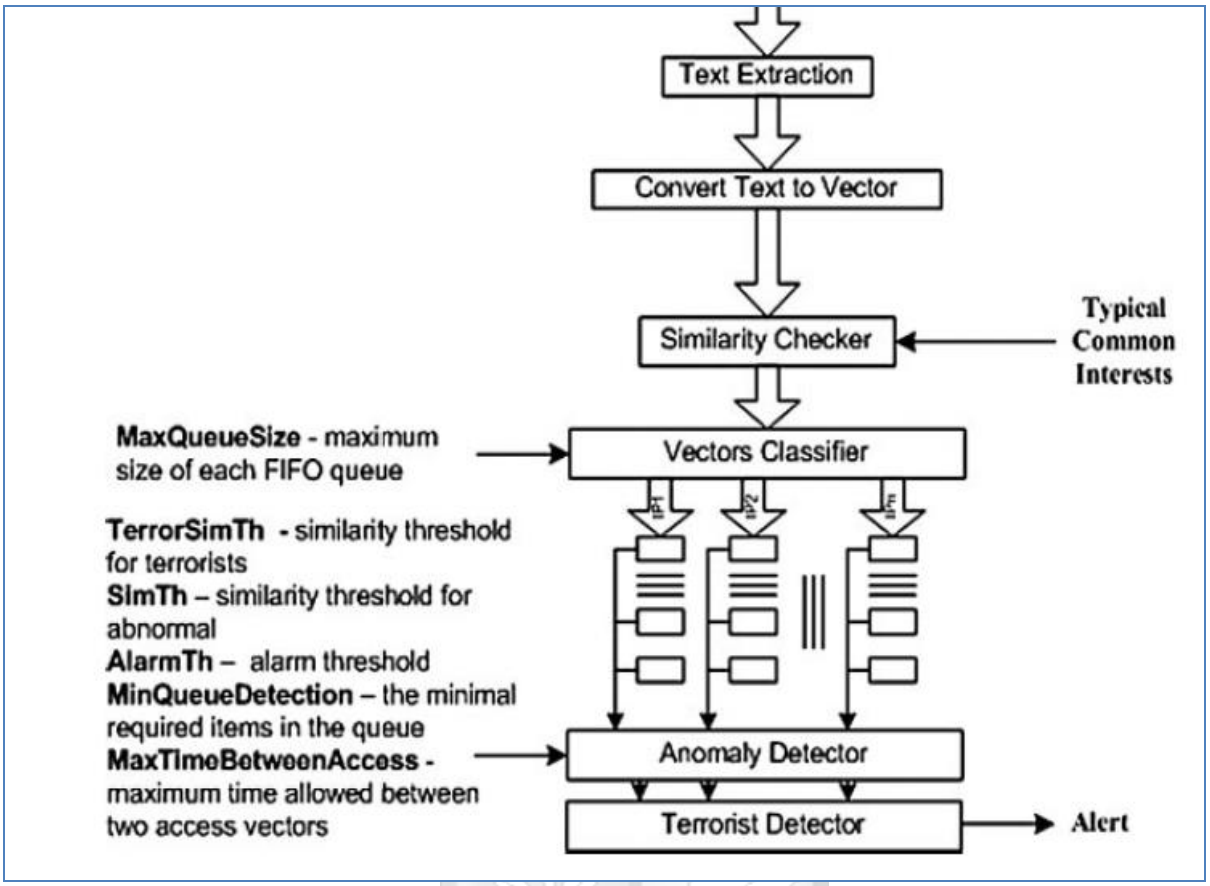

Figure 2.4: Detection mode

Once the learning mode (in figure 2.3) is completed, the system can switch to the detection mode. In the detection mode (in figure 2.4), the system detects users who are accessing typical content by intercepting HTML web pages. Each new incoming intercepted web page is converted into a vector of weighted terms and similarity is computed between the current vector and the known group profiles.

The detection algorithm makes a decision whether the user has recently visited the site or a series of pages by considering the history of a user's visit to the site. If in the learning mode, a profile of typical terrorist content has been built, then the system will also check the similarity of abnormal users to terrorist content. In the detection mode the pages that users access is captured, filtered, and transferred to a vector representation then the clustering module accesses the collected vectors and stores in the database and performs unsupervised clustering (e.g., using k-means algorithm) resulting in n clusters. The number of clusters Cn is one of the system parameters specified by the users (Yuval elovici et al, 2007).

2.7.2 Deviation detection

This method uses specific measures to study data that differs markedly from the rest of the data. An example of an application that uses this method is the outlier detection system. Outlier detection system identifies and finds patterns in data that do not conform to expected behavior. Outliers or anomalies can be detected without knowing the data set's distribution or needing any labelled training samples (Svetlana C, 2005). In the context of credit card fraud detection, a fraudulent transaction can be seen as an outlier which behaves differently comparing to legitimate transactions hence it can easily be spotted (Svetlana C, 2005).

Outliers sometimes called Anomalies are patterns in data that do not conform to a well-defined notion of normal behavior (Varun, Arindam, & Vipin, 2009). Figure 2.5 illustrates anomalies in a simple 2-dimensional data set. The data has two normal regions, N1 and N2, since most observations lie in these two regions. Points that are sufficiently far away from the regions, e.g., points o1 and o2, and points in region O3, are anomalies.

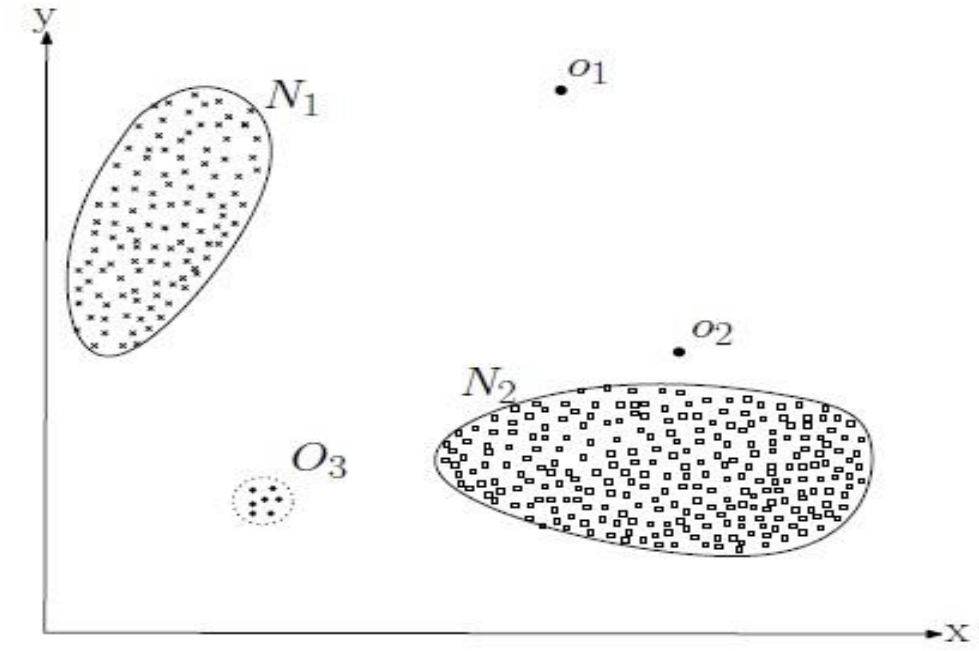

Figure 2.5: A simple example of anomalies in a 2-dimensional data set

Anomalies might be induced in the data for a variety of reasons, such as malicious activity, e.g., credit card fraud, cyber-intrusion, terrorist activity or breakdown of a system, but all of the reasons have a common characteristic that are interesting to the analyst (VARUN, ARINDAM, & VIPIN, 2009). Figure 2.6 shows the above mentioned key components associated with any anomaly detection technique.

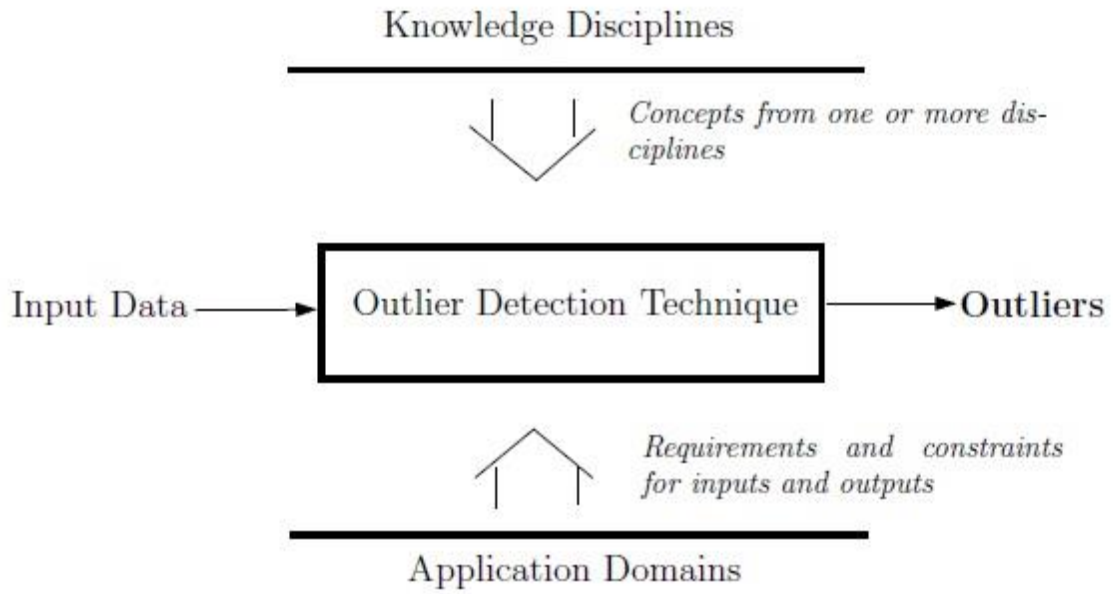

Figure 2.6: components of outliers' detection technique

This technique has extensive use in a wide variety of applications such as military surveillance for terrorist activities. The law enforcement officers can apply this method to detect fraud, crime analyses and network intrusion. However, some of these activities that this method guard against may appear to be normal making it difficult for this method to detect (Hsuchun et al, 2004).

2.7.3 Classification

Classification tries to find common features amongst different crime entities and organizes them into predefined classes. It is commonly used to predict crime patterns because it lessens time to identify crime entities. The investigators can use this method to reveal the identity of cyber criminals who use internet to spread radical information to the public. However, this technique requires a predefined classification scheme and a large amount training data and testing data because if there will be any missing data from the training set then accuracy of prediction would be limited thereby making it unsuitable for crime prediction (Hsuchun et al ,2004).

2.8 Common issues of sentiment analysis

2.8.1 Technical challenges

A number of technical challenges have been observed in sentiment analysis. A tweet may have sentiments expressed on different objects. In such a case, the problem lies in identifying the object on which a sentiment has been expressed without which the opinion is of little use (Liu, 2010). Secondly, many automated systems cannot differentiate between sarcasm and sincere text, or correctly analyze specific contextual meaning of a word. The use of acronyms like “lol” or word abbreviations also poses interpretation challenges (Meena & Joao, 2013).

2.8.2 Privacy concerns

The ease of communication has allowed the public to freely exchange anything they wish through the social media platform. It is difficult for a person to control dissemination of his/her personal information regarding his/her private life. Opinion mining involves the use of personal data of some kind, one of the most obvious ethical issues. This is a violation of people’s information privacy (James, 2012).

2.9 Conceptual model

The summary of the design and architecture of the model is given by the figure 2.7 below. The figure clearly shows the relationships between individual elements that make the model. Bag of words technique was used to develop the model.

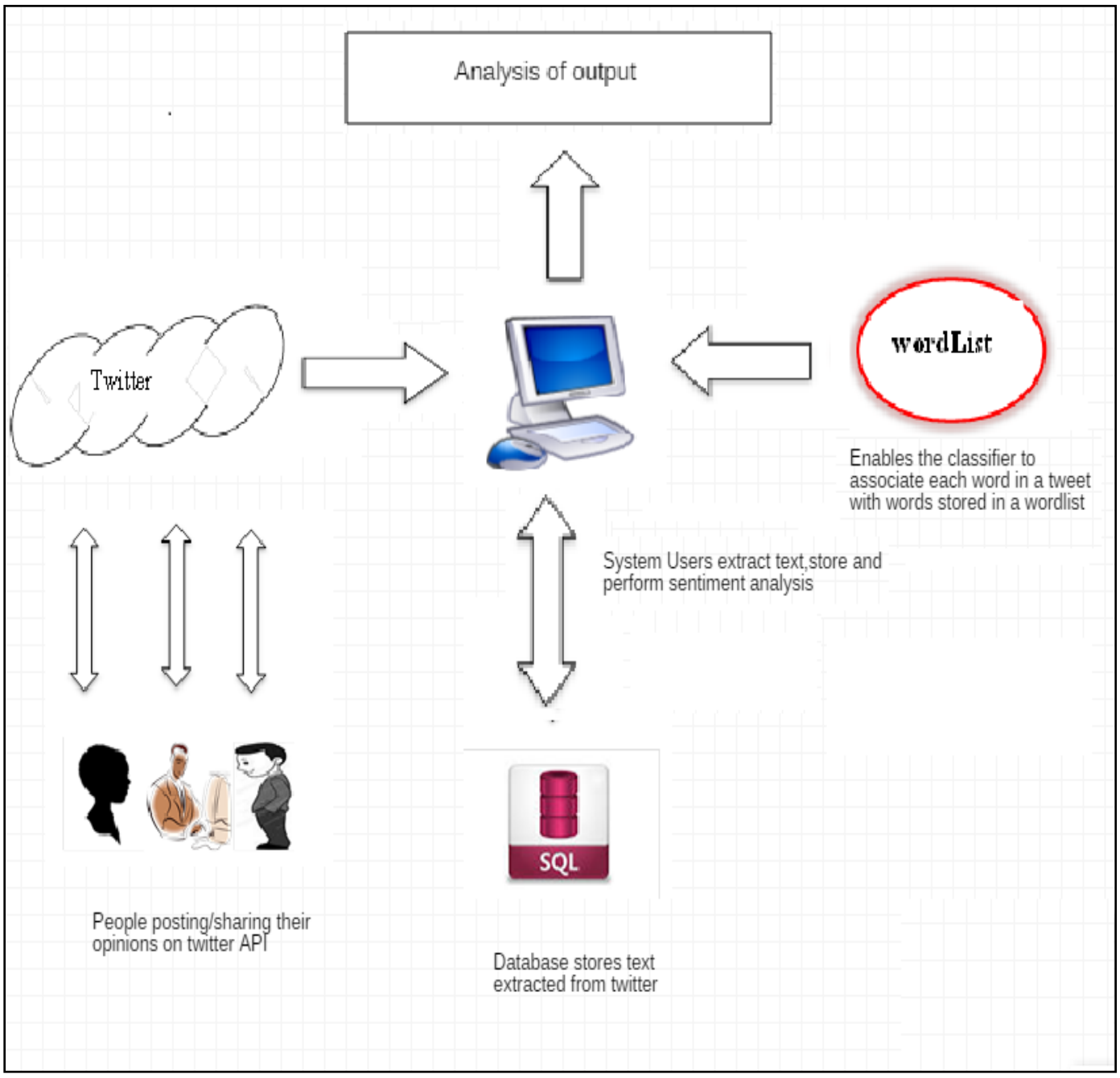

Figure 2.7: Conceptual model

CHAPTER THREE: METHODOLOGY

3.0 Introduction

Research methodology may be described as an orderly way of resolving a research problem. According to (Kothari, 2004) it can be termed as scientific research study. This chapter highlights the various methods and procedures that the researcher adopted in conducting the study. It describes the procedures that were used to collect and analyze data collected from twitter and finally develop the sentiment analyzer. The methodology covers the following major areas: requirement analysis, research design, data collection methods, implementation and verification.

3.1 Requirement analysis

Requirement analysis involved data collection from twitter streaming API using keywords for topics related to terrorists' activities. The streaming API allows access to global streaming of twitter data that is filtered using keywords. This process was repeated a number of times until a desired number of datasets were achieved. The data fetched from the site was stored in a SQLite3 database because SQLite3 database allows faster read/write operations and segmentation of data. The twitter data is stored according to the following headers: tweet id, date and time of tweet creation, actual text of the tweet, sender's place, latitude and longitude.

3.1.1 User requirements definition

Python programming language was used by the author to develop the model however the programming development tools were not passed to the users of the system. The users will only be given a report generated by the system on how sentiments can be used to track down terrorists from their hide out and locations. It will be upon them to make a decision of doing further investigation based on the report.

3.1.2 Requirements specification

The results generated by the system should be understood by users without any difficulties and should be in form of graphs and maps portraying the distribution of sentiments which can be easily visualized by the users. It is expected that the accuracy level of the classifier should be above 50% for it to be acceptable. This is so because computers do not achieve 100% accuracy

when classifying twitter data simply because the words used on twitter do not constitute formal language. They involve acronyms, emoticons and slang.

3.1.3 Functional requirements

Functional requirements define what is to be achieved by the system. The system should collect data from twitter according to specific topics related to terrorism within the Kenyan context and store the tweets in SQLite3 database for further processing.

3.2 Research design

The research design depended on requirement analysis and it was summarized as shown in the table 3.1 below.

Table 3.1: Summary of design activities

Objective	Activities	How objective was achieved
Design	Collect, clean and categorize training data from twitter	The twitter streaming API was used to collect data and stored in SQLite3 database then cleaned and stored in CSV files.
Develop	Design and build an opinion classifier which will categorize data as positive, negative and neutral	Bag of words approach and Python programming was used develop the classifier
Evaluation for efficacy	Customize a classification tool that focuses on the terrorist activities within Kenyan context	A python Scripts was developed to collect tweets and perform sentiment analysis and mapping of sentiment on a map

3.3 Data collection methods

These are techniques used for gathering data for research. The methods to be employed was determined by the type of research being conducted (Kothari, 2004). Data collection is very essential in research because without it the researcher cannot make factual and informed conclusions about the problem statement. The twitter streaming API was used to collect data and then stored in SQLite3 database for further processing.

3.3.1 Observation

Observational technique is employed to figure out social and organization requirements. Twitter data collected was examined to find out whether after classification the system correctly classified data into the three classes, namely negative, positive and neutral class.

3.3.2 Questionnaires

Questionnaires render an alternative to interviews and they are used to find out information about the system. They are composed of questions that are required about what is to be achieved by the system. The questionnaires are effective when sending similar information to a number of users. It also allows one to get an idea about the quality and performance of the system. The questions are structured and constitute an array of replies that beseech the users to tick one of them. It employed a mixture of open ended and closed ended questions.

3.3.3 Interviews

The interviews are structured with a predefined agenda that seek to ask questions that are focused to the problem statement and objectives of the system for quick and valid feedback**.**

3.4 Data Analysis

Ultimately the data collected, was analyzed using python libraries with the intention of formulating meaningful information from data collected from twitter. Data presentation was attained by use of graphs and maps.

3.5 Implementation of the system

The model was implemented using bag of words technique which uses Wordlist that is constructed based on hierarchical database model to give the correct scores with respect to the keywords. Python programming language was used to develop the model because it has many libraries and incorporate SQlite3 database in its libraries. SQlite3 database was used as database

management system (DBMS) for the model and used to store data collected from twitter streaming API. This was achieved by running a python Scripts created for data collection.

3.5.1. Bag of words model

Bag of words approach is a method of document classification where the frequency of each word is used as feature for training and developing a classifier. A text is represented as a bag (multi set) of words disregarding grammar and even words order but keeping multiplicity. It is similar to naives bayes classifier which work by correlating the use of tokens and using naves bayes theorem to calculate a probability of the absence or presence of a feature in a class (Pang & Lee, 2008).

Bag of words technique uses unigram features to carry out sentiment analysis and classification on twitter data. N-gram is technique of finding n-grams words from a given document (Pang & Lee, 2008). The commonly used model for this techniques includes unigram (n-1), bigrams (n=2) and trigrams (n=3). Unigram consists of all individual words present in the text; bigrams defines a pair adjacent words. Each pair words form a single bigram (Pang & Lee, 2008).

3.5.2 Unigram model

The system uses unigram technique to perform classification. This process involves building a database or CSV files called Wordlist which comprises of both positive and negative words. A tweet is then represented as a bag of words which is then broken into individual words. Each word is then matched to the words stored in the wordlist. When there is a match the counter is incremented or decremented by a fixed number depending on the weight or value assigned to each word in the wordlist. When this process is completed the classifier classifies the tweet as positive, negative or neutral.

CHAPTER FOUR: SYSTEM DESIGN AND ARCHITECTURE

4.0 Introduction

The various machine learning techniques reviewed assisted in selecting an appropriate platform for the development of the application. Use case and sequence diagrams were used to study the components of the system and how they interact with each other. It also shows the system architecture. The components that make up the system are; data collection, data cleansing, classification, sentiment analysis and mapping of sentiment.

4.1 System design

Systems design is the dissection of a system into its component pieces for purposes of studying how those component pieces interact and work (Gemino & Parker, 2009). The use case diagram in figure 4.2 shows the functional and non-functional requirements of the system. The components in the use case diagram are: twitter, which represent the twitter server, the user, and the end-user who is focused to use the information in the database.

4.1.1 Functional requirements

The functional and non-functional requirements are stated below;

1. Connected to twitter and fetched tweets based on keywords related to terrorism within the Kenyan context.
   a. No duplication of the fetched tweets.
   b. Retrieved the metadata of each tweet along with text content and coordinates (the longitude and the latitude). The tweets were stored in the database with the following headers;

i. Tweet Id.

ii. Date and time of tweet creation.

iii. Actual text of the tweet.

iv. Sender's place.

v. Geo coordinates (latitude and longitude).

c. Sent http request to twitter for every 30 seconds interval to fetch the tweets.

2. Created a database table to store tweet text along with extracted information of geo-coordinates.

3. Performed cleaning of tweets and stored in a database table.

4. Performed sentiment analysis on each tweet text and calculated polarity score as follows;

i. Positive.

ii. Negative.

iii. Neutral (zero).

5. Linked the classified tweet text corresponding to the geographic coordinates on to the map Geo-coding plays an important role in representing physical location on visual maps.

4.1.2 Non-functional requirements

i. Register the application with twitter and get the access keys.

ii. Provide consumer key, consumer secret and access tokens to Twitter streaming API to gain access to collect data.

iii. Configure global variables for SQlite3 Database.

4.1.3 Software requirements

i. Folium libraries – library for linking tweets to a map.

ii. Plotly libraries – library for printing graphs.

iii. Python 2.7 – open source software which provided the platform for the development of the model.

iv. Sqlite3 libraries – provided the database management system for the model.

v. Twitter API libraries – allowed the connection to twitter to extract tweets.

4.2 Use Case diagram

Figure 4.2 represents the use case diagram and the interaction between different components of the model and control flows. User is a main component who interacts with the twitter to get access based on search criteria. The user also depends on the following modules: data collection,

sentiment analysis and geocoding module. The user interacts with all the modules, collect data and store in the database. The sequence diagram is focused to facilitate a database of processed tweets fetched from the twitter. The sentiment analysis process as represented by the system involves the data collection, data cleansing, classification, sentiment analysis and mapping of sentiments. The results of this process are the information of processed data which is visualized in various forms like maps and graphs. To connect to twitter server, the model requires an internet connection, but the user can make use of data by using stand-alone applications.

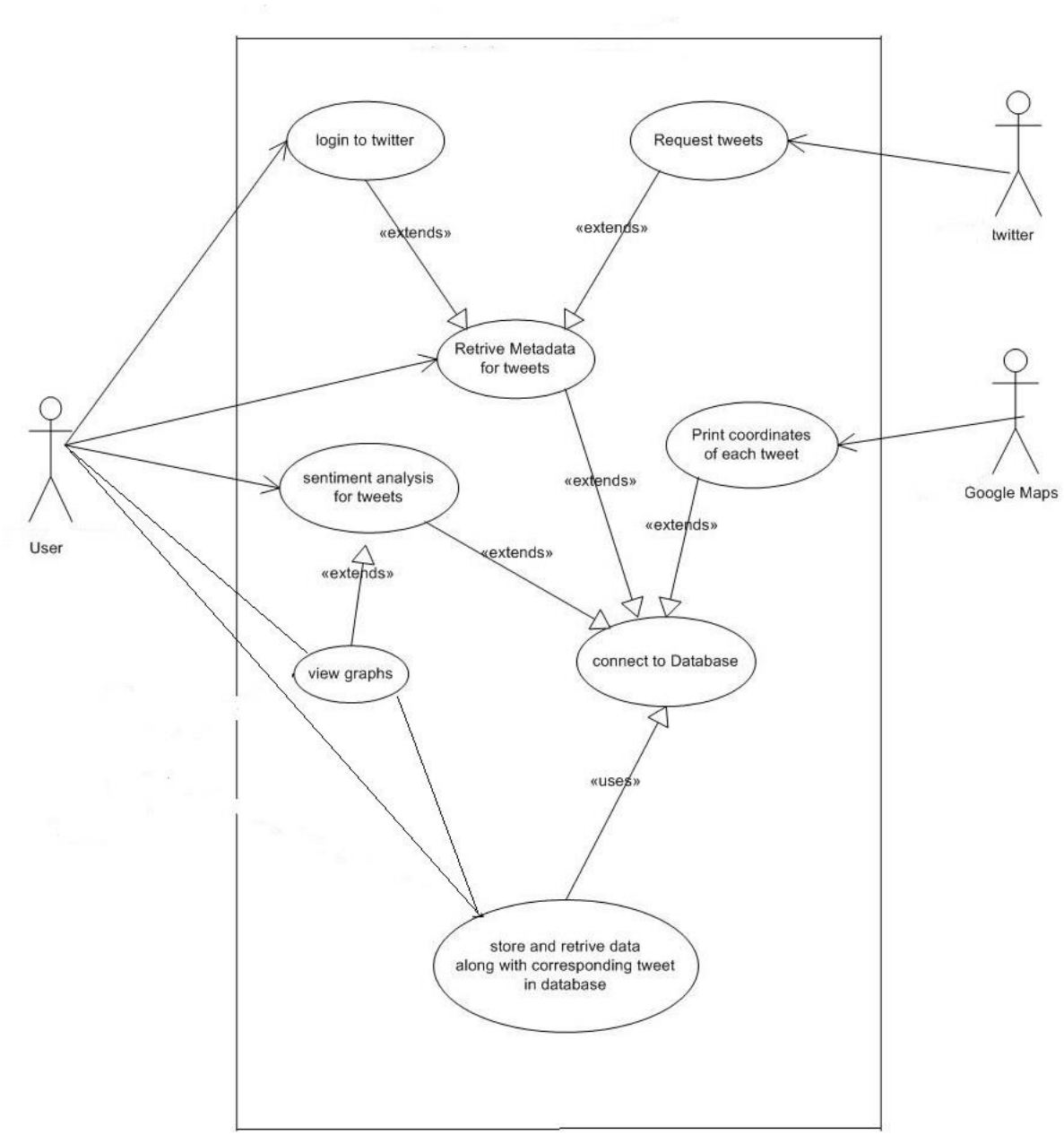

Figure 4.2: System components

4.3 Sequence diagram

Sequence diagrams envisage the interactive behavior of the system. It shows the behavior of the system, the message flow, the structural organization of the objects and the interaction among objects. The following sequence diagrams explain the operations which took part in the system.

4.3.1 Sequence diagram for data collection

This is the first stage in the system. The use case is activated when the user runs the python script. This is defined by figure 4.3.1.

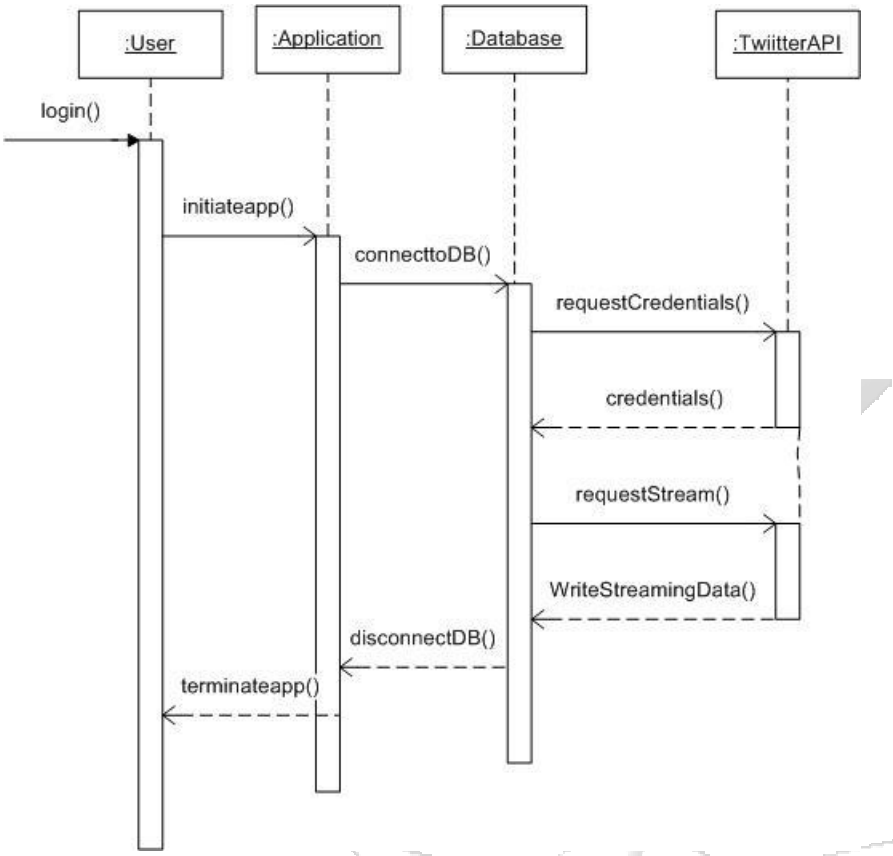

Figure 4.3.1: Data collection

4.3.1.1 Sequence of events

i. The process is activated by the user

ii. The connection to the database is established by the system

iii. System establishes connection with twitter Streaming API

iv. The credentials to access API is provided by the system

v. The streaming of data is initialized by the API

vi. The data is written into the database by the system

vii. The user terminates the system

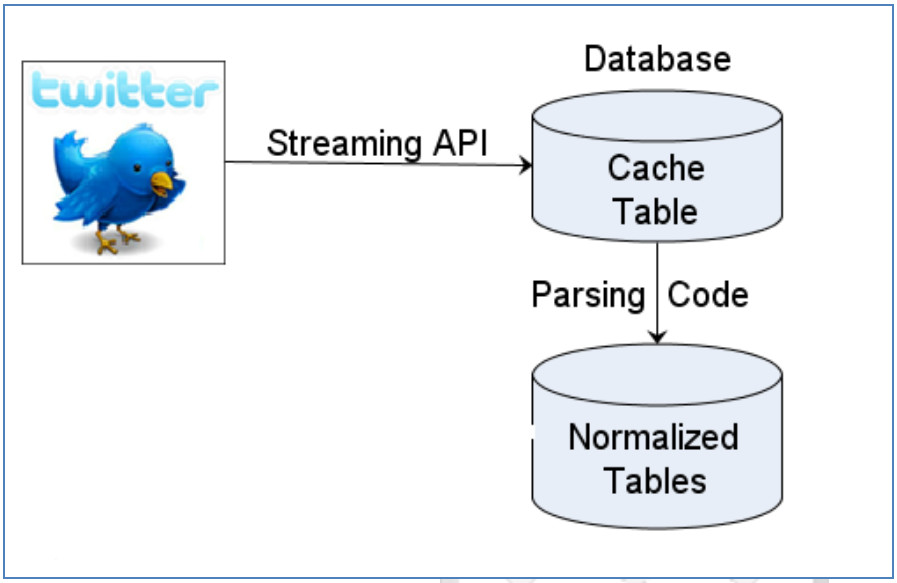

Figure 4.3.2: Streaming process

The streaming process collect the input tweets and carry out filtering before storing the tweets to a database as shown in figure 4.3.2. The tweets are stored in the database table according to the following column headers; Tweet Id, Date and time of tweet creation, Actual text of the tweet, Sender's place, latitude and longitude. The hypertext link transfer protocol (HTTP) handling process, queries the web for results in response to user requests.

4.3.2 Sequence diagram for data cleansing

It is a process whereby unwanted contents in the data collected are removed. If unwanted contents are left in the database, it will be very difficult to classify raw data accurately. The condition for this stage is that data collection has to be completed successfully before data cleansing begins. A python script is then executed to activate the system. The sequence diagram for data cleansing is shown in figure 4.3.3.

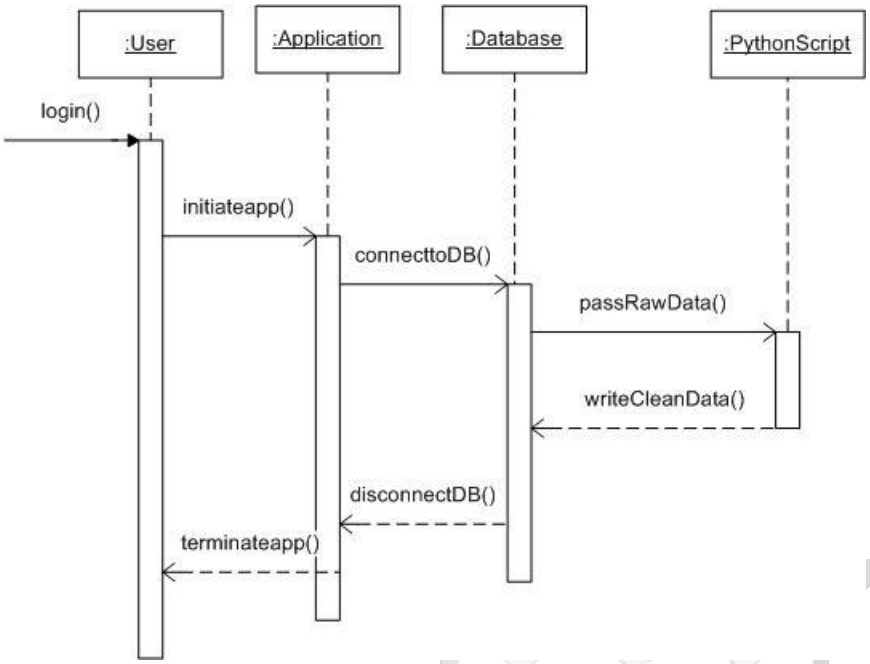

Figure 4.3.3: Data cleansing

4.3.2.1 Sequence of events

i. The process is activated by the user

ii. System establishes connection with database

iii. Data is read by the system

iv. Unwanted contents are removed by system

v. The cleaned data is written back to database by the system

vi. The system then terminates

4.3.3 Sequence diagram for data classification

This stage focuses on classification of cleansed tweets stored in CSV file into any of the three classes (positive, negative, or neutral). The preconditions for this stage are that data collection and data cleansing have to be completed successfully before classification begins. A python script is then executed to activate the use case. The use case for classification is shown in figure 4.3.4.

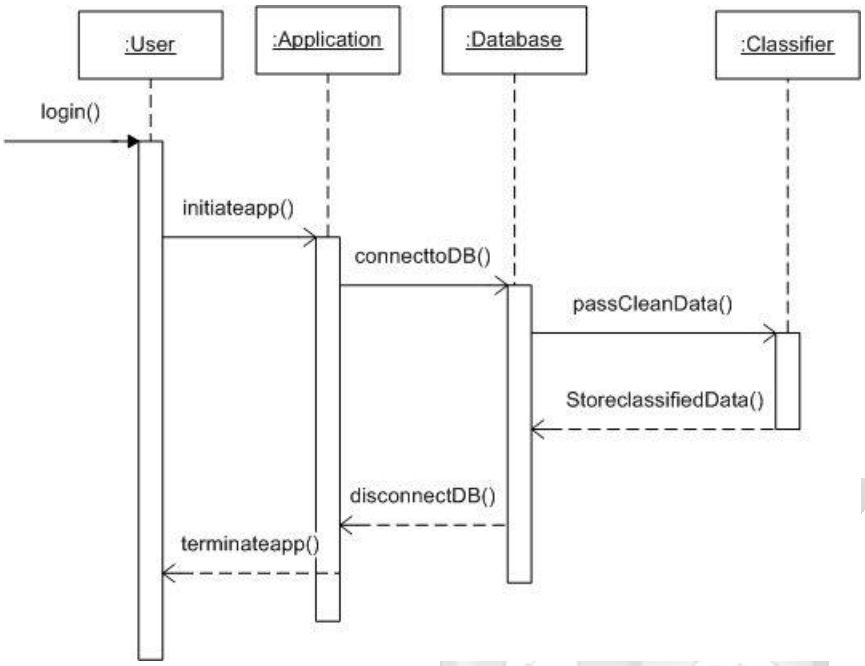

Figure 4.3.4: Data classification

4.3.3.1 Sequence of events

i. The process is activated by the user

ii. System establishes connection with database

iii. Data is read by the system

iv. Data is classified into classes by the system

v. The system then terminates

4.3.4 Sequence diagram for sentiment analysis

The output of the analysis produced is represented using a graph so that the results can be visualized by the user. The preconditions for this stage are that data collection, data cleansing and classification have to be completed successfully before analysis of data begins. A python script is then executed to activate the use case. The use case for analysis is shown in figure 4.3.5.

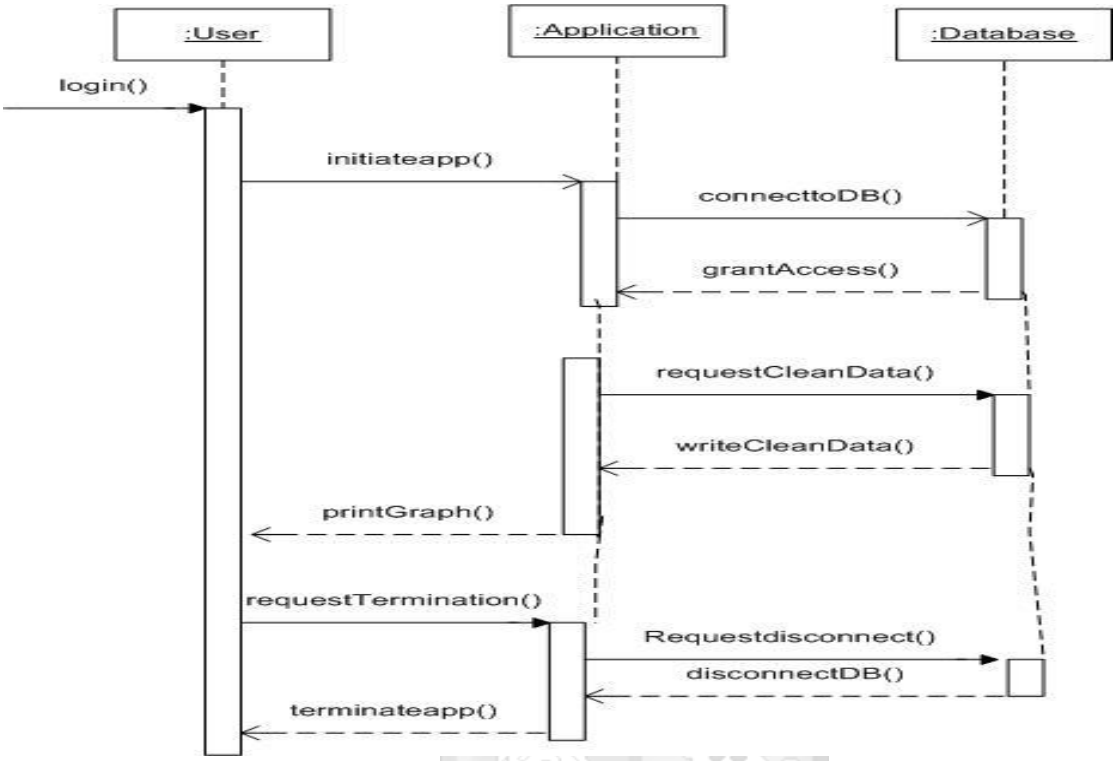

Figure 4.3.5: Data analysis

4.3.4.1 Sequence of events

i. The process is activated by the user

ii. System establishes connection with database

iii. Data is read by the system

iv. The data is analyzed by system

v. System print a graph showing (positive, negative and neutral sentiments)

vi. The system then terminates

4.3.5 Sequence diagram for sentiment mapping

This is a process of mapping sentiments on to a map using their coordinates. The map is produced with markers distributed across various points on the map so that the sentiments can be easily visualized by the user. The data collection, data cleansing and classification have to be completed successfully before sentiment mapping begins. A python script is then executed to activate the use case. The use case for sentiment mapping is shown in figure 4.3.6.

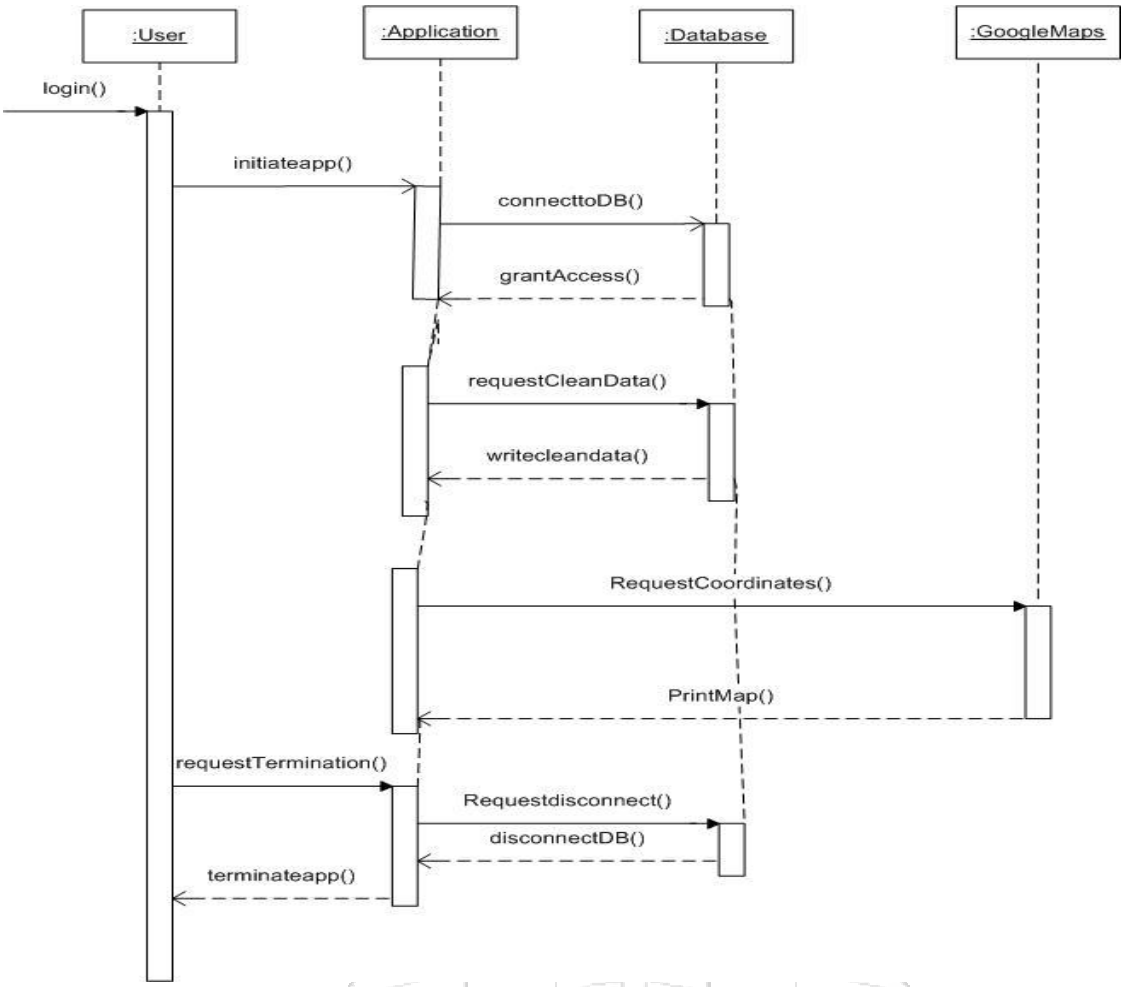

Figure 4.3.6: Sentiment mapping

4.3.5.1 Sequence of events

i. The process is activated by the user

ii. System establishes connection with database

iii. Data is read by the system

iv. System links Sentiments to coordinates on the map

v. System print a map showing (distribution of sentiments)

vi. The system then terminates

4.4 System architecture

The figure 4.4 shows the components that make up the system. The components present the tasks that a user can carry out using the system. These tasks are; data collection, data cleansing, classification, analysis and mapping of sentiment onto a map.

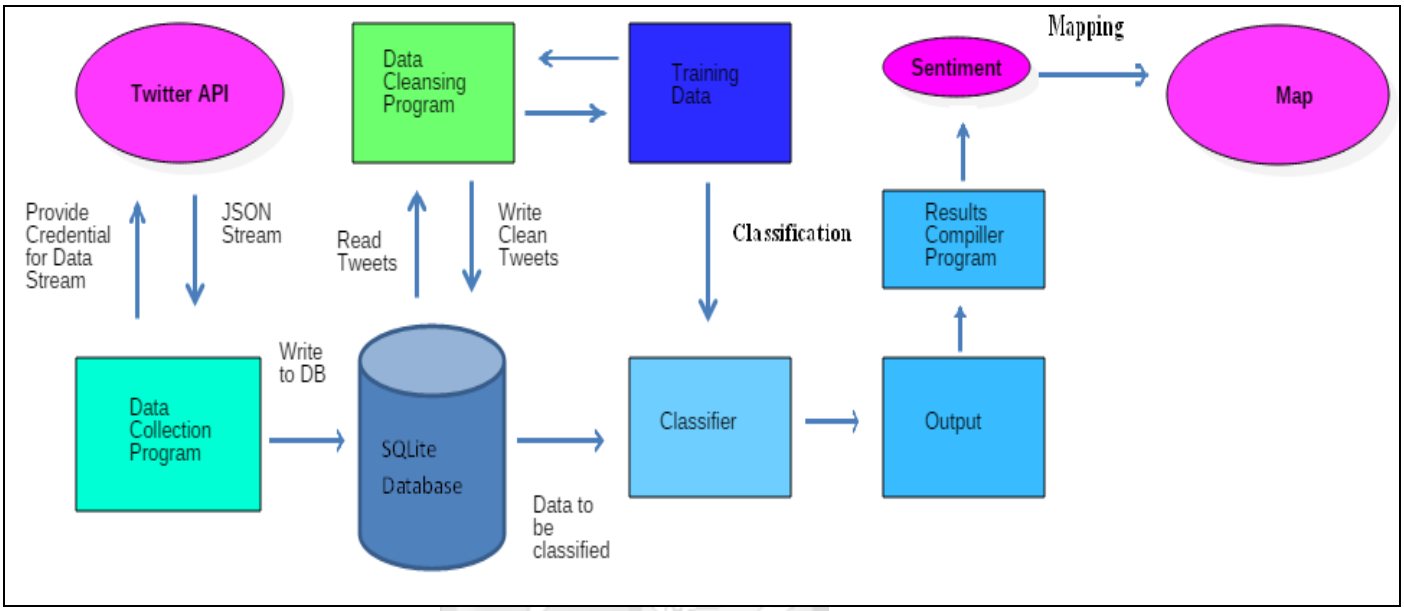

Figure 4.4: Structure of the system

4.5 System analysis

This model performs like client-server system in which the user (client) communicates with twitter server while accessing tweets. The system communicates through internet satisfying client credential of the twitter server.

4.5.1 Data collection

The data collection was achieved using the public streams API. The public streams API is an efficient way of gathering information for purposes of data mining as it allows access to global streaming of twitter data that could be filtered using keywords. Python interface library was installed to interface with twitter's API.

Twitter has numerous regulations and rate limits imposed on its API such that the requests made for fetching data are limited to 30 seconds beyond which the fetch operations is time out and disconnected thus limiting the amount of data being collected. It is a requirement that all users must register an account and provide authentication details when fetching data from the twitter API. Figure 4.5.1 shows the authentication keys.

```
1 consumer_key = "4cAcTX56kxfCK8mTf7AG36HfV"
2 consumer_secret =
  "VUCtRfhzttoQRz3BJfws1jKXGyoqTyojaE0yWsf3VUTFTuX2qi"
3 access_token_key = "1535707669-
  fm2DjmAtsxCmjoBb7NOTI7NVKgTi3vUUbZKLusw"
4 access_token_secret =
  "GOszWV2VNu7a0WdR9mjLG0M8qPRx1PmeLNE9zCYpkRSe2"
```

Figure 4.5.1: Authentication keys

A python script with authentication details was created and executed to provide the connection to twitter API. This connection initialized the streaming process where data was pulled from twitter API and in a format of JSON and stored in SQLite3 database. The SQLite3 database was created with table called “tweets” which had a structure consisting of the following fields: id, date, tweet, location, latitude and longitude as shown in figure 4.5.2. Figure 4.5.3 and table 4.1 show an example of data collected in a JSON format. The python script that performed this action is shown in the appendix.

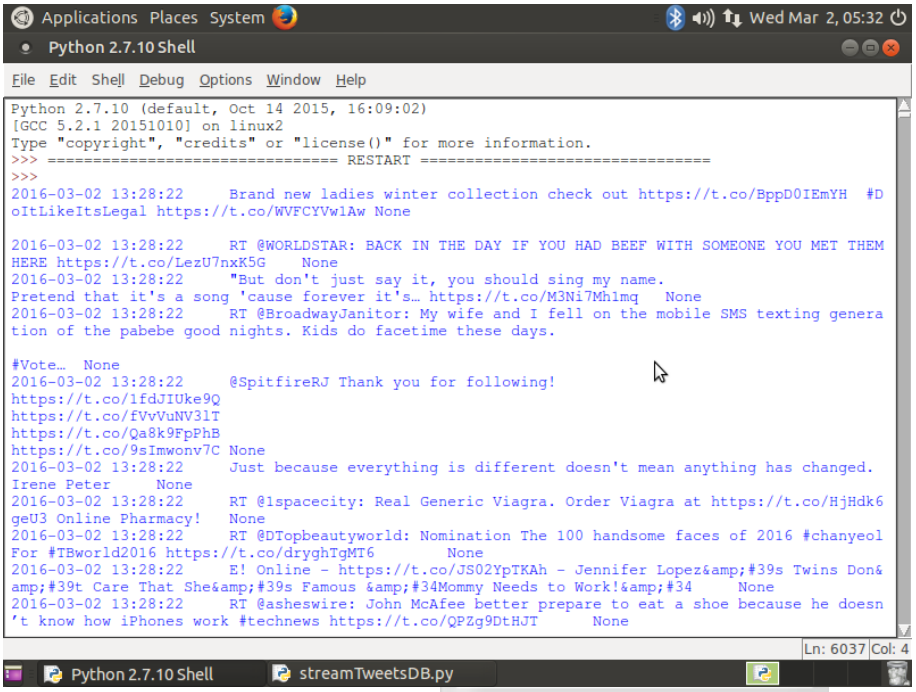

Figure 4.5.2: Sample of data streaming

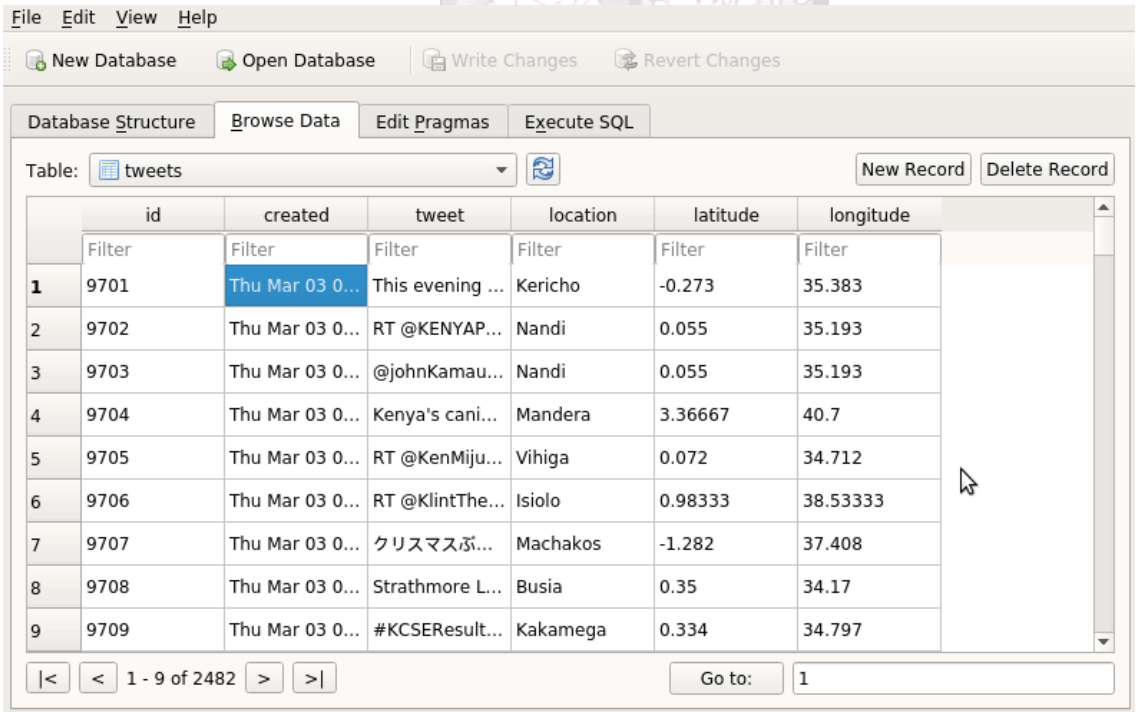

Figure 4.5.3: A sample of stored raw data

Table 4.1: Sample of collected data

ID	Date	Tweet	Location	Latitude	Longitude
12112	Tue Mar 08 16:23:23 +0000 2016	US Bombs al-shabaab Terror Camps Graduation ;killing over 150 Terrorists .Https://t.coKn7AMJFrNF	Embu	-0.425	37.531
12119	Tue Mar 08 16:23:30 +0000 2016	RT@BonifceMwangi :President @UKenyatta so you know ,Kenya Government paid Ksh 2.4 Billion for GSU Land and it has been grabbed @JBoinet	Kisumu	-0.069	34.64
12182	Tue Mar 08 16:23:46 +0000 2016	Terror fight neeeds multifacet Strategy .Https.//t.co/LPwLiMiY6P	Kakamega	0.334	34.797
12100	Tue Mar 08 16:23:04 +0000 2016	Video:Elephant rescued from a well in Kenya .Https./t.c/m^TuZKOxiH	Narok	-1.24076	35.7356

4.5.2 Training data

This is a dataset consisting of training examples and their corresponding target. In a supervised learning environment, the algorithms use these datasets to learn to map training examples to their corresponding targets. If the training process is correctly implemented, the algorithm should be able to generalize the training data so that it can map new data correctly **(**Pang & Lee, 2008). For this model, the training example is our sanitized data collected during data collection phase of this system. Table 4.2 and figure 4.5.4 show samples of training data.

Table 4.2: Sample of training data

ID	Tweet	Polarity	Value
12100	video:elephant rescued from a well in kenya .url	Positive	2
12112	us bombs al-shabaab terror camps graduation ;killing over 150 terrorists url	Negative	-2
12119	atuser :president atuser so you know ,kenya government paid ksh 2.4 billion for gsu land and it has been grabbed atuser	Neutral	0
12182	terror fight neeeds multifacet strategy url	Negative	-1

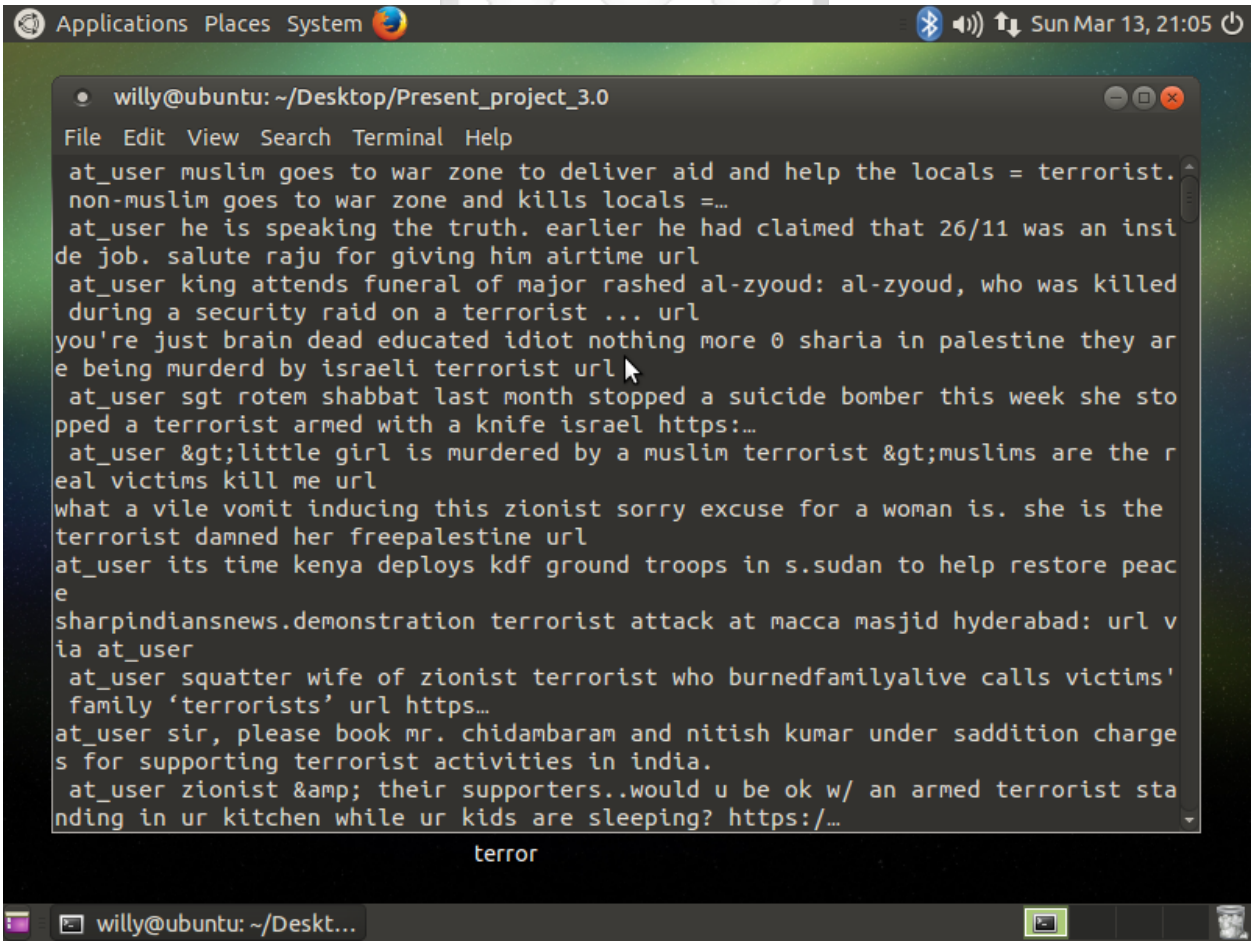

Figure 4.5.4: A Sample of training data

4.5.3 Data cleansing

It is a process whereby unwanted content are removed from the training data which is used as the input tweets. The unwanted content is any word within the tweet that is not useful for classification. The unwanted contents are shown in table 4.3 and table 4.4 whereas clean tweets are shown in figure 4.5.5. The python script that performed this action is shown in the appendix.

Table 4.3: Sample of unwanted content and action

S/No.	Unwanted Conten	Wanted Content	Action
1.	#word	Word	Replaced
2.	@username	AT_USER	Converted
3.	https://	URL	Converted
4.	Additional White Space ‘ ‘	“ ”	Removed
5.	Retweet	RT	Removed
6.	Uppercase	lowercase	Converted

Table 4.4: Sample unwanted content and cleaned tweets

ID	Date	Tweet	Location	Latitude	Longitude	Type
12112	Tue Mar 08 16:23:23 +0000 2016	US Bombs al-shabaab Terror Camps Graduation ;killing over 150 Terrorists .Https://t.coKn7AMJFrNF	Embu	-0.425	37.531	Raw
		us bombs al-shabaab terror camps graduation ;killing over 150 terrorists url				**Clean**
12119	Tue Mar 08 16:23:30	RT@BonifceMwangi :President @UKenyatta	Kisumu	-0.069	34.64	Raw

	+0000 2016	so you know ,Kenya Government paid Ksh 2.4 Billion for GSU Land and it has been grabbed @JBoinet				
		atuser :president atuser so you know ,kenya government paid ksh 2.4 billion for gsu land and it has been grabbed atuser				**Clean**
12182	Tue Mar 08 16:23:46 +0000 2016	Terror fight neeeds multifacet Strategy .Https.//t.co/LPwLiMiY6P				Raw
		terror fight neeeds multifacet strategy url				**Clean**
12100	Tue Mar 08 16:23:04 +0000 2016	Video:Elephant rescued from a well in Kenya .Https./t.c/m^TuZKOxiH	Narok	-1.24076	35.7356	Raw
		video:elephant rescued from a well in kenya .url				**Clean**

Each tweet was read from the database and processed in order to clean the undesirable data. This is shown in figure 4.5.5. A python script that performed this action is shown in the appendix.

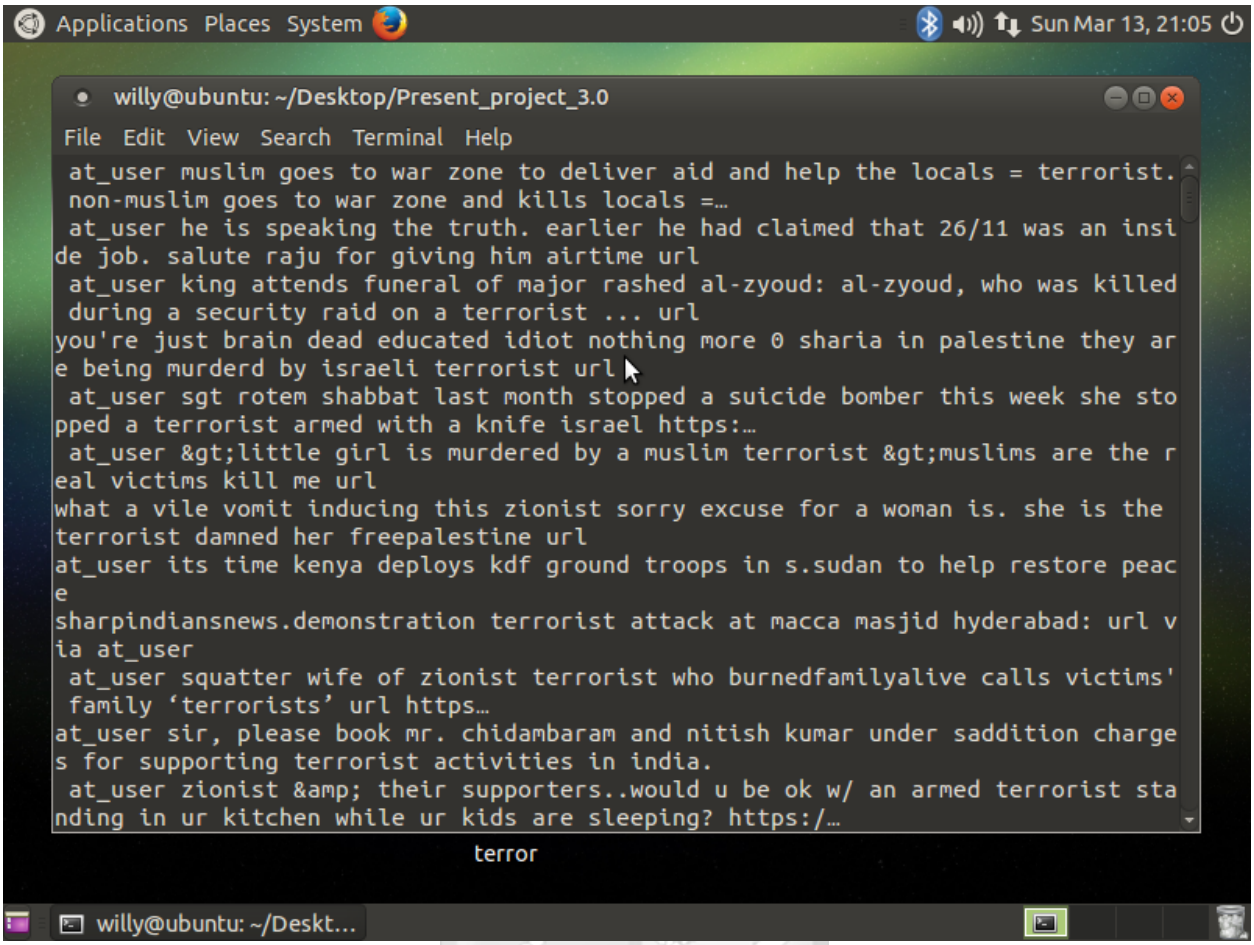

Figure 4.5.5: A Sample of clean tweets

4.5.4 Word list

This is a list of negative and positive words created and stored in CSV file. The classifier read this file and associates each tweet with it such that if there are many positive words in a tweet than negative words it classifies the tweet as positive and vice versa. The wordlist is shown in figure 4.5.6 and figure 4.5.7. The corpus in the "wordlist" was generated from the www.thesauras.com dictionary. This process involved searching of the synonyms and antonyms of words relating to terrorism and stored in a database labeled negative whereas none terrorist words were stored in a database labeled positive. The dictionary contained 2,084 positive words and 4,269 negative words. The model computes the total score of each text by calculating the aggregate score of words in a source text.

Figure 4.5.6: A sample of negative wordlist

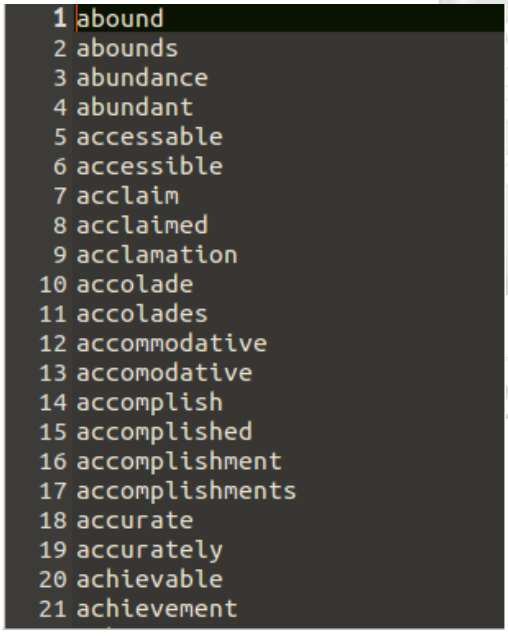

Figure 4.5.7: A sample positive wordlist

4.5.5 Data Classification

This is a process whereby data collected from Twitter is classified into three classes' namely negative, positive and neutral class. A classifier was created to read tweets from the database which involved building of a database or CSV files of both positive and negative wordlists. A tweet was then represented as a bag of words which was broken down into individual words. Each word was matched to the words stored in the positive and negative wordlist. When there was a match the counter was incremented or decremented by a fixed number depending on the weight or value assigned to each word in the wordlist. When this process was completed the classifier classified the tweet as positive, negative or neutral. When the classification was completed, the results were saved in a text file as shown in the appendix. An example of classified tweets is shown in table 4.5 and figure 4.5.8 below. A python script that performed this action is shown in the appendix.

Table 4.5: Sample of classified tweet

ID	**Tweet**	**Polarity**	**Value**
12100	video:elephant rescued from a well in kenya .url	Positive	2
12112	us bombs al-shabaab terror camps graduation ;killing over 150 terrorists url	Negative	-2
12119	atuser :president atuser so you know ,kenya government paid ksh 2.4 billion for gsu land and it has been grabbed atuser	Neutral	0
12182	terror fight neeeds multifacet strategy url	Negative	-1

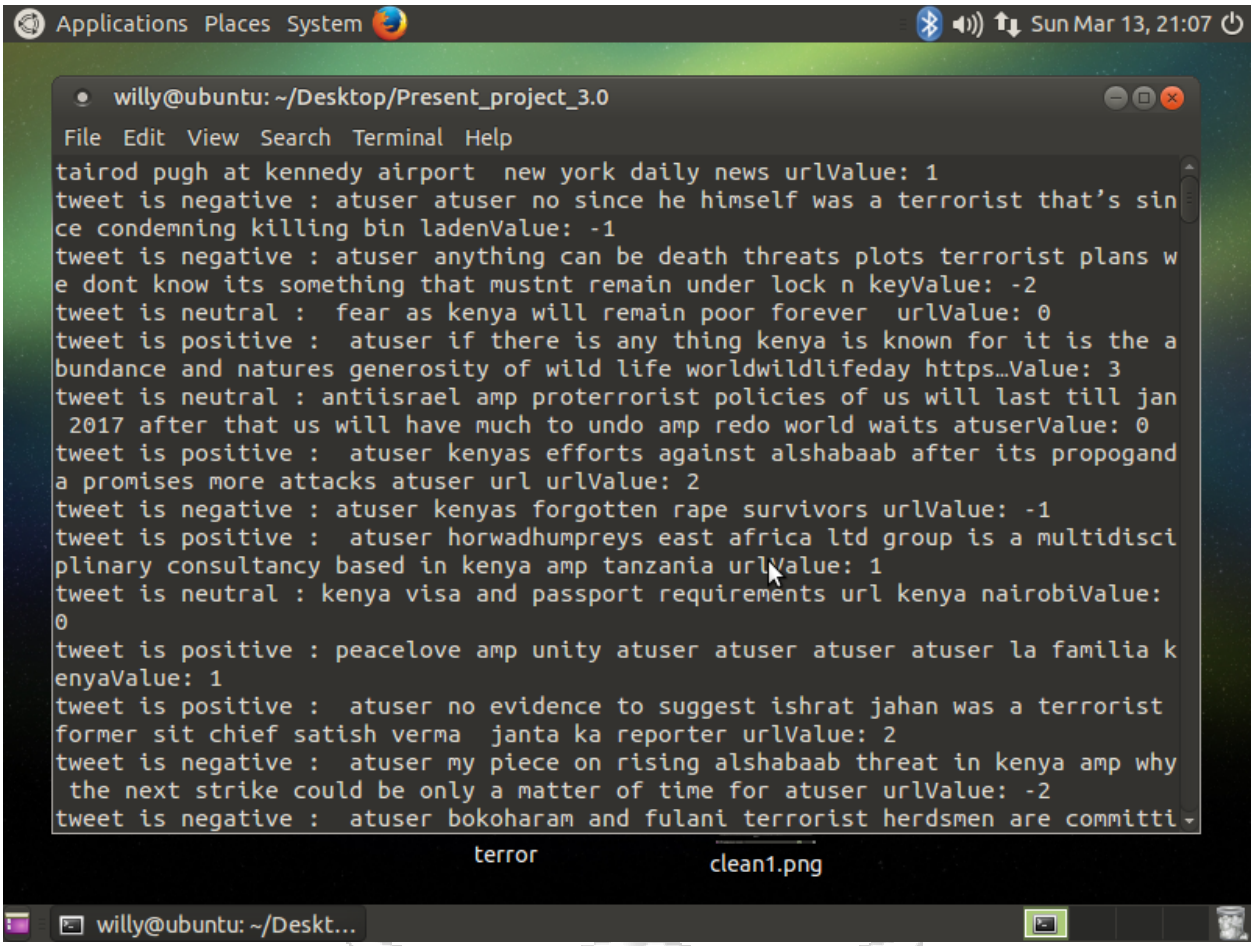

Figure 4.5.8: A sample of classified data

The results of the classification included the coordinates which specified the location of the tweeter. The coordinates help in mapping the sentiment on to the map as shown in the figure 4.5.9.

```
 1 Positive Tweet,Kericho,-0.273,35.383,2
 2 Neutral Tweet,Nandi,0.055,35.193,0
 3 Negative Tweet,Nandi,0.055,35.193,-1
 4 Negative Tweet,Mandera,3.36667,40.7,-1
 5 Neutral Tweet,Vihiga,0.072,34.712,0
 6 Positive Tweet,Isiolo,0.98333,38.53333,1
 7 Neutral Tweet,Machakos,-1.282,37.408,0
 8 Neutral Tweet,Busia,0.35,34.17,0
 9 Neutral Tweet,Kakamega,0.334,34.797,0
10 Positive Tweet,Wajir,1.75,40.01667,1
11 Positive Tweet,Marsabit,2.96667,37.6,2
12 Positive Tweet,Isiolo,0.98333,38.53333,1
13 Neutral Tweet,Taita Taveta,-3.4,38.37,0
14 Neutral Tweet,Samburu,1.33333,37.11667,0
15 Positive Tweet,West Pokot,1.75,35.25,1
16 Neutral Tweet,Elegeyo-Marakwet,0.99,35.55,0
17 Positive Tweet,Homa Bay,-0.666,34.481,3
18 Neutral Tweet,Siaya,0.105,34.302,0
19 Neutral Tweet,Garissa,-0.172,40.041,0
20 Positive Tweet,Nairobi,-1.28333,36.83333,2
21 Neutral Tweet,Isiolo,0.98333,38.53333,0
```

Figure 4.5.9: A sample of classified data with locations

Python script was created to perform the calculation of percentage of positive, neural and negative tweets stored in a CSV file as shown in the appendix. Figure 4.5.10 and 4.5.11 shows analysis of tweets and a bar graph of number of tweets against their polarities respectively. The results of classification are shown in table 4.6.

```
willy@ubuntu:~/Desktop/Present_project_3.0$ python plot_polarity.py
20
275
51
willy@ubuntu:~/Desktop/Present_project_3.0$
(process:2292): GLib-CRITICAL **: g_slice_set_config: assertion 'sys_page_size =
= 0' failed

(process:2301): GLib-CRITICAL **: g_slice_set_config: assertion 'sys_page_size =
= 0' failed
```

Figure 4.5.10: Analysis of classes

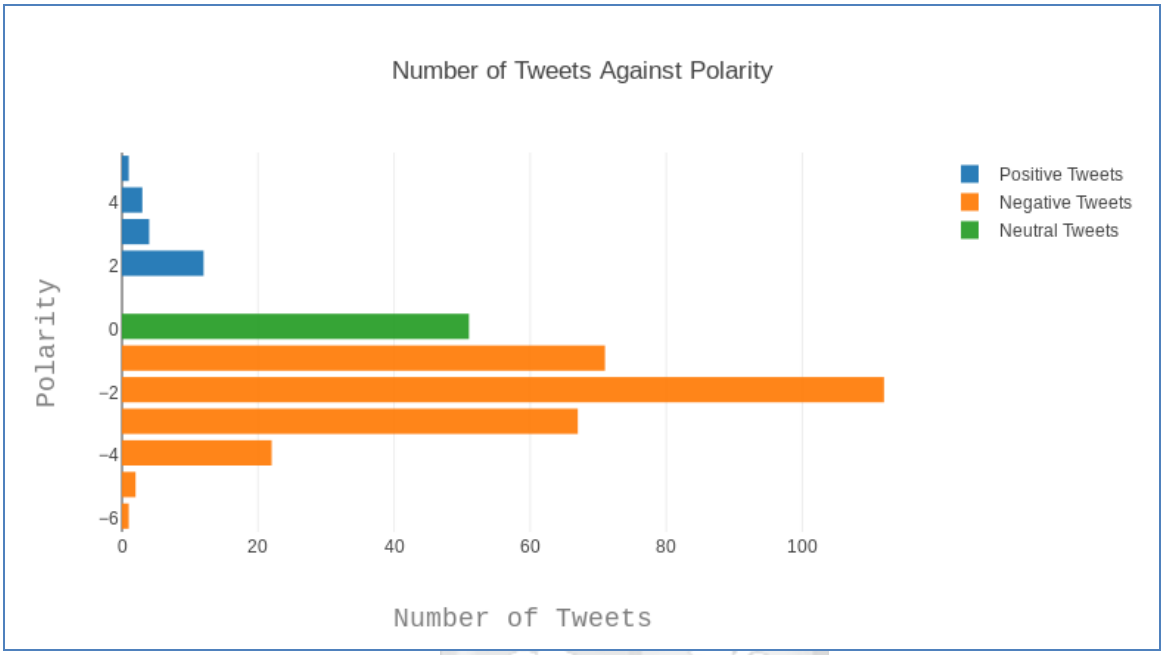

Figure 4.5.11: Sentiment polarity graph

Table 4.6: Results

S/No.	Instances	Total
1.	Total Number of instances	**346**
2.	Total instances of positive class	20
3.	Total instances of negative class	275
4.	Total instances of neutral class	51

4.3.6 Sentiment Mapping

This is a process of mapping sentiments on to the map. The map is produced with markers distributed across various points on it so that the sentiments can be easily visualized by the user. The three classes are indicated on the map using markers with different colors i.e. red for negative, green for positive and yellow for neutral sentiment. The maps produced by the system are shown in figure 4.5.12, 4.5.13, 4.5.15 and 4.5.15 below. The python script that performed this action is shown in the appendix.

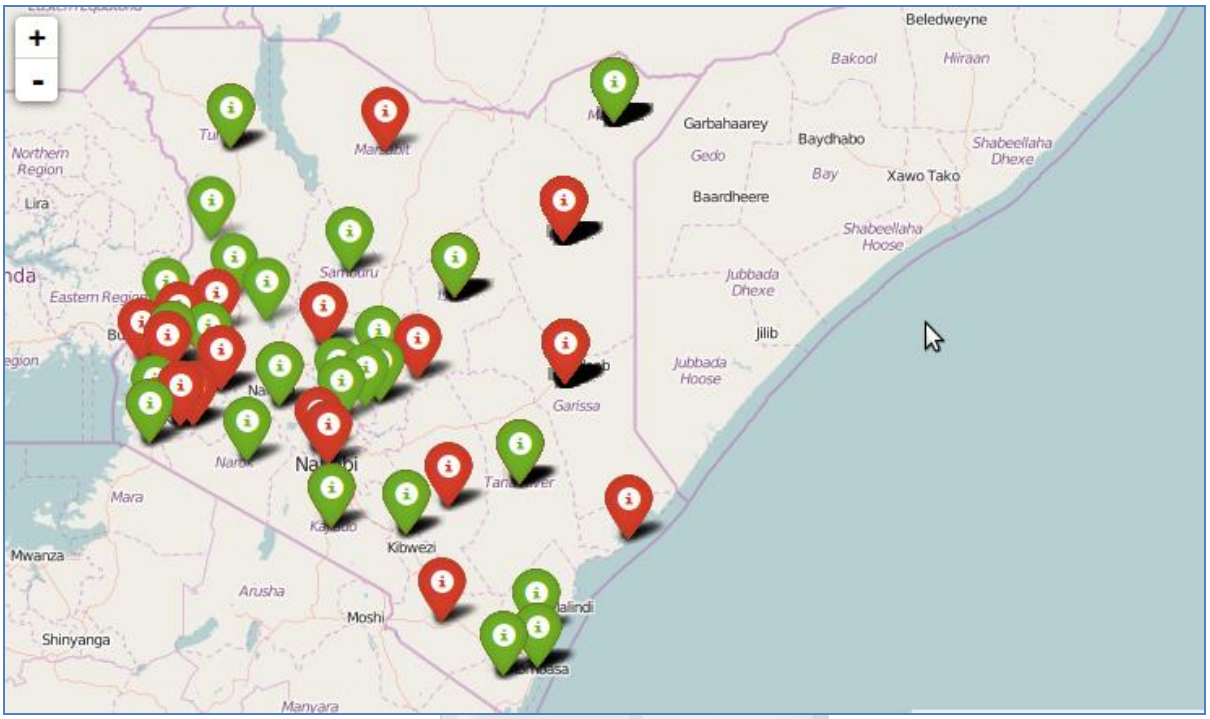

Figure 4.5.12: Sentiment distribution map 1

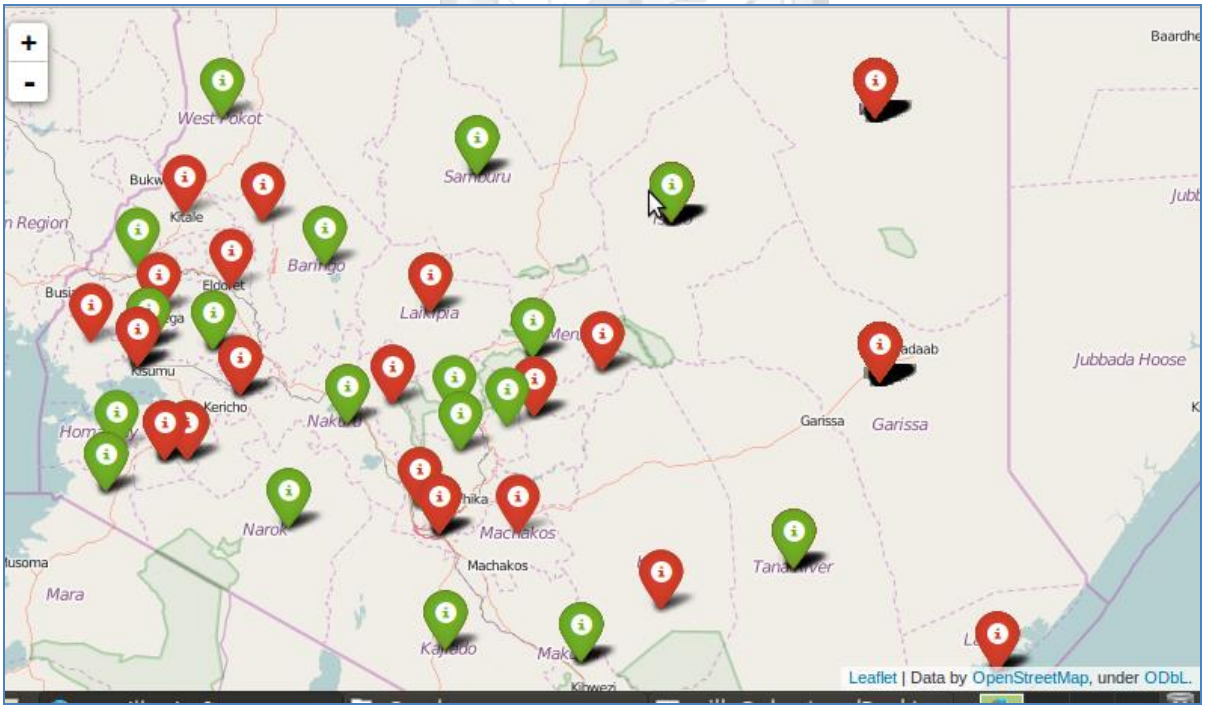

Figure 4.5.13: Sentiment distribution map 2

The map below shows the exact location of the tweeter

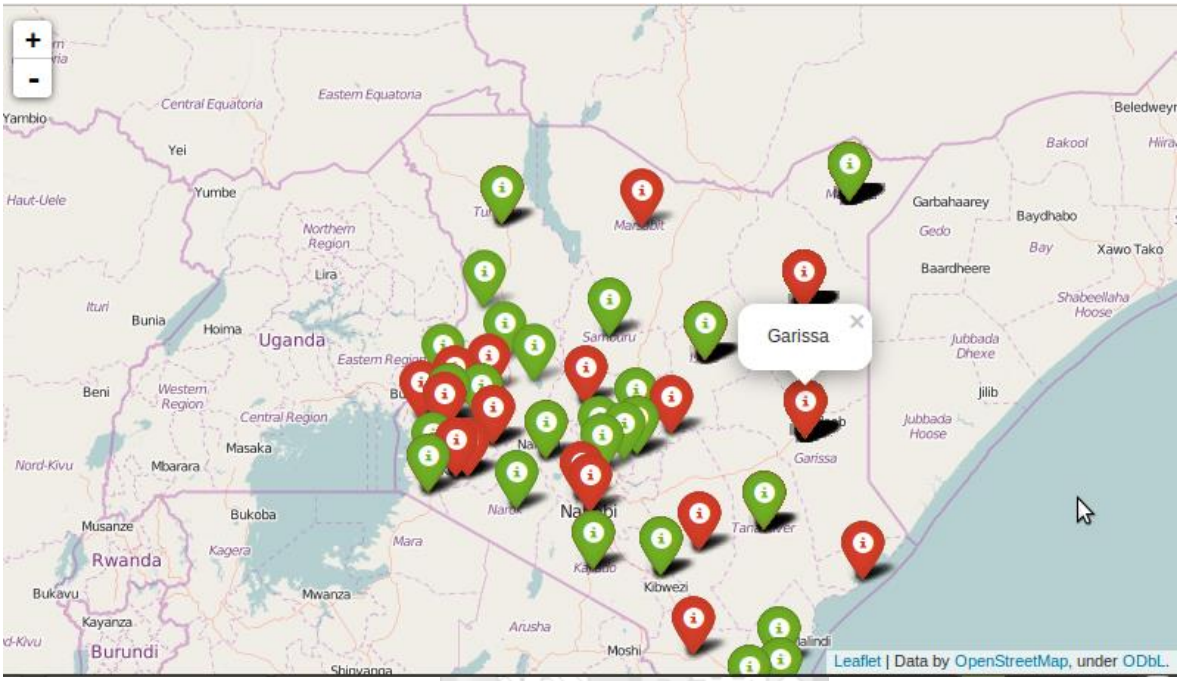

Figure 4.5.14: Location of the tweeter 1

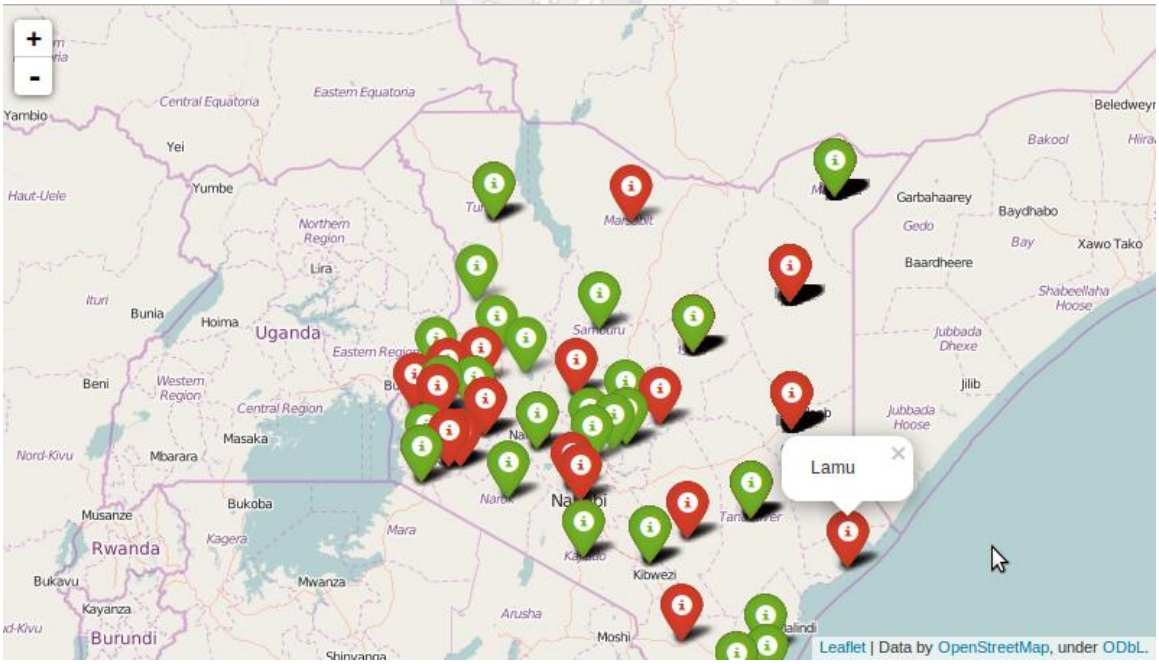

Figure 4.5.15: Location of the tweeter2

CHAPTER FIVE: IMPLEMENTATION AND TESTING

5.0 Introduction

The implementation of the system involved the following: data mining, text processing and machine learning techniques. Testing of the system was given to a group of IT professionals and intelligent analysts, to test the system, its functionality, and usability in accordance with the application objectives.

5.1 Implementation of the system

It outlines a brief description of the platform that was used to implement the model. The platform entailed the principle libraries, environments and main software languages used in the development of the model.

5.2 Platform

A computer system with 64bit Intel i3 Processor (4GB RAM) running a Windows Operating System and virtual machine running Linux operating system was used to run the model which carried out the task of data collection, data cleansing, classification, sentiment analysis and mapping of sentiment on to a map.

5.2.1 Python

Python 2.7.6 was used to implement this model because it is an open source software, has many libraries and a command interpreter which allows users write code, test and debug codes quickly as there is no compilation step required. Python has an inbuilt data types for strings, lists, and more which make it a versatile and robust high level programming language.

5.2.2 SQLite3

SQLte3 database is open source software and is incorporated in python libraries. It was used as database management system (DBMS) for this model because it allows faster read/write operations and segmentation of the data to be stored and support multithreaded applications. The SQLite3 database is shown in figure 5.1.

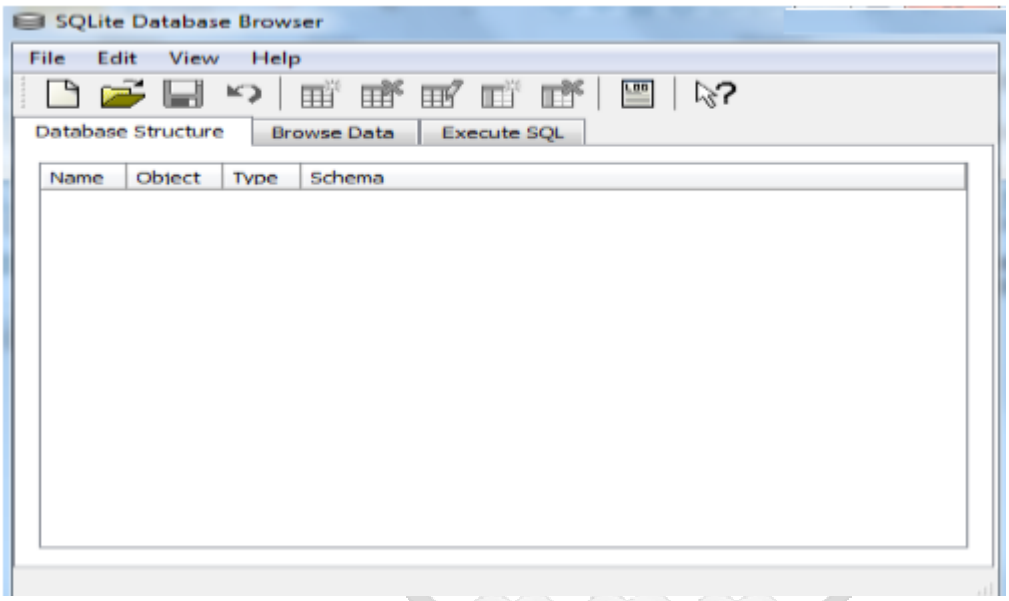

Figure 5.1: SQLite3 database

5.2.3 CSV (Comma Separated Values)

CSV files were also used to store the information in a comma separated values (CSV) form.

5.2.4 Bag of words technique and n-grams algorithm

The model uses bag of words technique which is based on a word weight instead of the term frequency of each word as in the standard BOW. The standard BOW model uses a huge lexicon which has duplications of word and repetition. This lexicon is built manually which requires to create a positive and negative words list by recognize the sentiment polarities based on the personal observation (Doaa, 2016). This approach takes a big time and efforts to compute the total score of sentiments of twitter data. Another problem of BOW is low accuracy because the standard BOW model neglects text grammatically and ordering of words (Doaa, 2016).

The model has a small lexicon that reduces the standard lexicon of BOW and deal with adjectives, nouns, verbs, adverbs, adjectives, prefixes, suffixes or other grammatical classes as a word. The lexicon was constructed automatically and was based on hierarchical database model to give the correct scores with respect a topic features and keywords. The lexical approaches save time and hasten searching process for each word (Doaa, 2016). After some pre-processing, the dictionary contained 2,084 positive words and 4,269 negative words. The model computes the total score of each text by calculating the aggregate score of words in a source text.

5.3 Testing

Ease of use and clarity of the system is tested to ensure that the system meets the requirements of the users.

5.3.1 Test results

The test results from the classifier when evaluated using a dataset containing 346 tweets collected from twitter streaming API is displayed in table 5.1.

Table 5.1: Overall sentiment

S/No.	Instances	Total
1.	Total Number of instances	346
2.	Total instances of positive class	20
3.	Total instances of negative class	275
4.	Total instances of neutral class	51
5.	Percentage of instances classified positive	5.80%
6.	Percentage of instances classified negative	79.50%
7.	Percentage of instances classified neutral	14.70%

5.3.2 Visualization

The visualization of the sentiment collected from twitter API is displayed and presented by the map in figure 5.2.

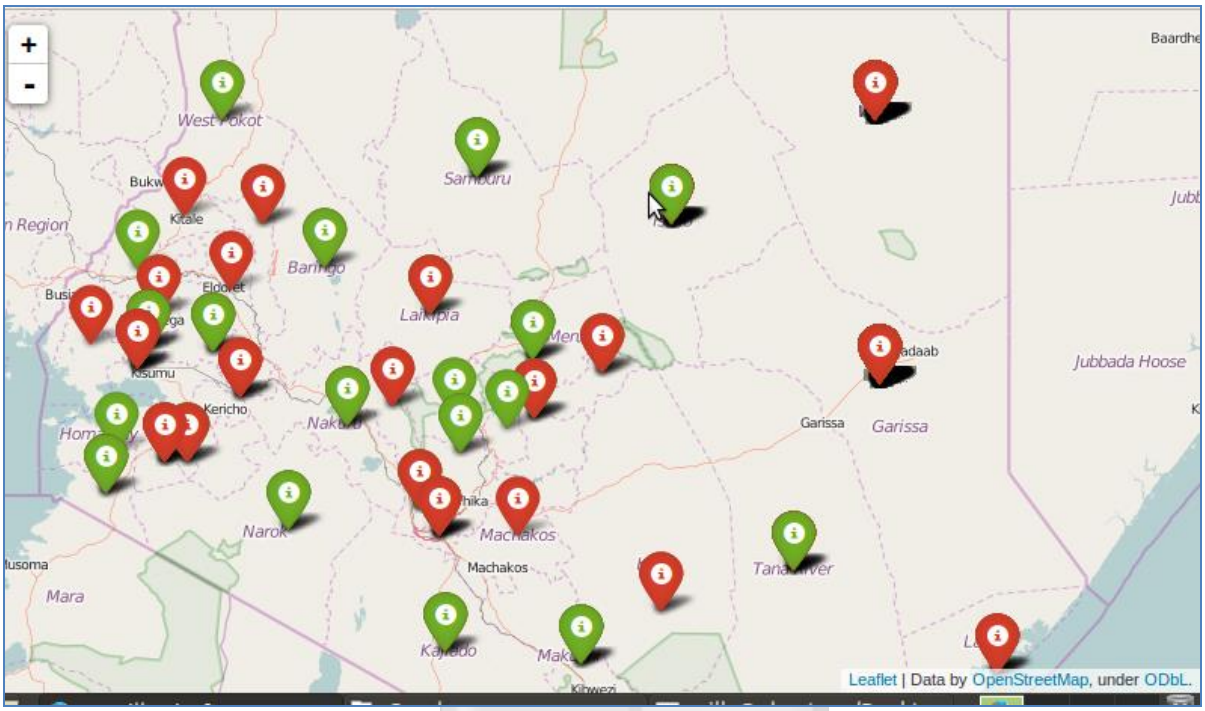

Figure 5.2: Sentiment distribution

The results show that there was high percentage of negative tweets related to terrorism compared to positive tweets. The locations where the sentiments originated were indicated on the map using markers i.e. red for negative, green for positive and yellow for neutral sentiment as shown in figure 5.2. These results can be used by law enforcement officers to give them investigative leads and information that will help them in disrupting, exposing and uncovering terrorists' networks and their structure effectively which would otherwise take hours to uncover manually.

5.3.3 Evaluation of the model

The performance of sentiment classification can be measured by using the following equations; Accuracy = (TP+TN)/(TP+TN+FP+FN), Precision = TP/(TP+FP) and Recall = TP/(TP+FN). In which TP, FN, FP and TN refer respectively to the number of true positive instances, the number of false negative instances, the number of false positive instances and the number of true negative instances, as defined in the table 5.2 (Doaa,2016).

Table 5.2 Confusion Matrix

	Predicted Positives	**Predicted Negatives**
Actual Positive	TP	FN
Actual Negative	FP	TN

The true positive (TP) rate or recall is the rate at which positive text are predicted to be positive (R), whereas the true negative rate is the rate at which negative text are predicted to be negative. The accuracy represents the rate at which the model predicts results correctly (A) (Doaa, 2016). The precision also called the positive predictive rate, calculates how close the measured values are to each other (P) (Doaa, 2016).

Table 5.3: Overall system performance

S/No.	**Metric**	**Accuracy Score**
1.	Classification Accuracy	73%
2.	Precision	60
3.	Recall	15%

Table 5.3 shows a summary of the overall system performance; the accuracy level of the model was 73%, the rate at which positive text are predicted sometimes called recall was (15%) and the precision was (60%). The system performance in table 5.3 shows that the classifier was able to perform sentiment analysis with a classification accuracy of 73% thus making it more efficient to be used in classifying real–time twitter data. The distribution of sentiments on the map as indicated by markers represented the patterns and trends of terrorist activities in Kenya thus can be used by law enforcement officers to give them investigative leads and information that will help them in disrupting, exposing and uncovering terrorists' networks and their structure effectively.

CHAPTER SIX: CONCLUSIONS AND RECOMMENDATIONS

6.0 Introduction

This chapter reviews and summarizes the research study. Various techniques and methods written by different authors on sentiment analysis were researched and the technological options that were available for this model was identified. The technical, functional requirements and the design and architecture of the model were also defined. The bag of words technique was used for its implementation. Testing and evaluation of the model was done throughout its development.

6.1 Conclusions

The research was guided by the five objectives. First was, to determine the current terrorism detection methods in use; to the extent of this objective, indeed there are terrorism detection measures employed by the government though they do not adequately address the growing extremist minority and terrorist threat in Kenya. The second objective was to identify terrorist activities using the current methods; a number of limitations were documented in the earlier section of this document. The third objective was to review existing data mining techniques available for use in crime detection; a number of techniques were identified and their merit and demerits were analyzed in depth and documented. The fourth and final objective was to develop and test the system; the system was developed and the results and performance of the system were documented in one of the section of the research study.

6.2 Recommendations

This research has shown that it is possible to have an automated system that can expose and uncover terrorists' networks and their structure effectively which would otherwise take hours to uncover manually. However, this research was carried out in a small scale. The Detection of terrorists' activities is not only limited to terrorists' networks but it can be extended to identity other criminal activities within Kenya. Variations of words which keep on changing must be added into the wordlist. The systems accuracy can be enhanced by having more rules that covers these words.

6.3 Limitations

Some of the limitations of the research study were; first, the requests made for fetching data were limited to only 30 seconds beyond which the fetch operations were timed out and disconnected thus limited the amount of data collected. Lastly, there was a lot of slang language in the tweets; this made it difficult to prepare the training data and classification.

6.4 Future Work

Some of the future improvements that can be made to this model are; the model can be enhanced to be accessible on handheld devices such as mobile phones. This will help to mitigate the threat of terrorist activities by identify, exposing and uncovering their networks and structure quickly.

REFERENCES

Agel (2013), a framework for employees' appraisals based on inductive logic programming and data mining methods.

Amnesty International (AI) (2013), police reforms in Kenya a drop in the Ocean.

Avinash (2015), decision trees how to construct them and how to use them for classifying new data, Purdue University.

Bolla, Raja & Ashok (2014), crime pattern detection using online social media.

Brittany Justine & Patrick (2013), Westgate Tweets: A detailed Study in Information Forensics.

Bo & Lillian (2008), opinion mining and sentiment analysis, foundations and trends in Information Retrieval Vol. 2.

Chetan &Atul (2014), a scalable, lexicon based technique for sentiment analysis, international journal in foundations of computer science &technology (IJFCST) Vol.4.

Dingding, Shenghuo & Tao (2012), Sum View: A web-based engine for summarizing product reviews and customer opinions.

Doaa (2016), Enhancement Bag-of-Words Model for Solving the Challenges of Sentiment Analysis, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 7.

Fredrick (2013), Online Social Networks and Terrorism 2.0 in Developing Countries, International journal and computers science and networks solutions Volume 1.

Govindarajan & Romina (2013), a survey of classification methods and applications for sentiment analysis, the international journal of engineering and science (IJES).

International Association of Crime Analysts (IACA) (2008), Exploring Crime Analysis.

International Centre for Prevention of Crime (ICPC) (2008), police and crime prevention in Africa.

James (2012), social media and intelligence gathering.

Johnson,Shukla & Shukla (2012), on classifying the political sentiment of tweets.

Jytte (2015), tweeting the jihad, social media networks of western foreign fighters in Syria and iraq.

Kester, Quist-aphetsi & Mieee (2013), visualization and analysis of geographical crime patterns using formal concept analysis, international journal of remote sensing & geoscience (IJRSG), volume 2.

Kothari (2004), research methodology,methods and techniques. new delhi, new age international.

Leah K & Abdallah B (2009), Social Policy, Development and Governance in Kenya.

Meena & Joao (2013), marketing research: the role of sentiment analysis.

National Crime Research Centre (NCRC) (2012), center summary of a study on organized criminal gangs in Kenya.

Pang & Lee (2008), opinions mining and sentiment analysis. Foundations and trends in information extraction, vol 2.

Pete B et al (2014), tweeting the terror, modeling the social media reaction to the woolwich terrorist attack.

Pravesh & Mohd (2014), methodological study of opinion mining and sentiment analysis techniques, international journal on soft computing (IJSC) vol. 5.

Ravendra (2014), a proposed novel approach for sentiment analysis and opinion mining,iju, vol 5.

RuiXia,ChengqingZong & Shoushan (2011), ensemble of feature sets and classification algorithms for sentiment classification.

Samuel (2013), neglecting history and geopolitics in approaches to counterterrorism, african journal of criminology and justice studies (AJCJS) vol.7.

Svetlana C (2005), Outlier Detection in Clustering.

Tawunrat& Jeremy (2013), affect analysis of radical contents on web forums using sentiWordNet, international journal of innovation, management and technology, Vol. 4.

United Nation (UN) (2012), the use of the internet for terrorist purposes.

Varun, Arindam, & Vipin (2009), anomaly detection, ACM computing surveys,

Vishal & Sonawane (2016), sentiment analysis of twitter data: a survey of techniques, international journal of computer applications, volume 139.

Yuval elovici et al (2007), detection of access to terror-related web sites using an advanced terror detection system.

Weimann (2014), New Terrorism and New Media. Washington, DC: Commons Lab of the Woodrow Wilson International Center for Scholars.

APPENDIX

Python Programs

Sqlite3 database creation program

```
conn=sqlite3.connect('Terror2016.DB')
c=conn.cursor()

#c.execute ("CREATE TABLE TWEET (created TEXT,tweet TEXT,coordinates TEXT)")
```

Script for saving results

```
# Read tweets cleaned by File : cleanTweets.py
tweetsfile = open("twitter_data/cleanedtweets.csv")

# File to write final sentiment analysis
finalfile = open('twitter_data/last_tweet_sentiments.csv', 'wb')
writer = csv.writer(finalfile, delimiter=',')
```

Data collection program

```
    def on_status(self, status):

        try:
            tcreated = status.created_at
            ttext = status.text
            tcoordinates = status.coordinates

            print "%s\t%s\t%s" % (tcreated, ttext, tcoordinates)

            cur.execute("INSERT INTO TWEETS(created, tweet, coordinates) VALUES
            con.commit()



        except Exception, e:
            print >> sys.stderr, 'Encountered Exception:', e
            pass

    def on_error(self, status_code):
        print >> sys.stderr, 'Encountered error with status code:', status_code
        return True # Don't kill the stream

    def on_timeout(self):
        print >> sys.stderr, 'Timeout...'
        return True # Don't kill the stream

streaming_api = tweepy.streaming.Stream(auth, CustomStreamListener(), timeout=60


streaming_api.filter(track=['TERROR-kenya','alshabaab-KENYA'], languages=['en'])
```

Data cleaning program

```
#Convert www.* or https?://* to URL
tweet = re.sub('((www\.[^\s]+)|(https?://[^\s]+))','URL',tweet)

#Convert @username to AT_USER
tweet = re.sub('@[^\s]+','AT_USER',tweet)

#Remove additional white spaces
tweet = re.sub('[\s]+', ' ', tweet)

#Remove retweets "RT"
tweet=re.sub('RT','',tweet)

#Replace #word with word
tweet = re.sub(r'#([^\s]+)', r'\1', tweet)

#trim
tweet = tweet.strip('\'"')

#Convert to lower case
tweet = tweet.lower()

return tweet
```

Classifier

```
# Increment the counters on each word found in negative and postive bag of w
for word in words:
    if word in positive_words and word != ' ':
        positive_counter=positive_counter+1

    elif word in negative_words and word != ' ':
        negative_counter=negative_counter+1
processedTweet = tweet_processed
# When postive words are more that negative, append positive counter
if positive_counter > negative_counter:
    positive_counts.append(positive_counter)
    print "tweet is positive : "+ tweet_processed + "Value: " + str(positive
    holder = "Positive Tweet"
    writer.writerow([holder,item[3],item[4],item[5],positive_counter])

# When negative words are more that postive, append negative counter
elif negative_counter > positive_counter:
    negative_counter = negative_counter * -1
    negative_counts.append(negative_counter)
    print "tweet is negative : "+ tweet_processed + "Value: " + str(negative
    holder = "Negative Tweet"
    writer.writerow([holder,item[3],item[4],item[5],negative_counter])
```

Sentiment visualization program

```
# For every line in the CSV file, append respective list
for line in readdata:
    if int(line) > 0:
        positivetweets.append(line)
    elif int(line) < 0:
        negativetweets.append(line)
    elif int(line) == 0:
        neutraltweets.append(line)

print len(positivetweets)
print len(negativetweets)
print len(neutraltweets)


############## Scatter Plot

# map our data (in lists) to a Scatter plot
positivetweetsPlot = go.Scatter(y = positivetweets, opacity = 0.95, name = "Posi
negativetweetsPlot = go.Scatter(y = negativetweets, opacity = 0.95, name = "Nega
neutraltweetsPlot = go.Scatter(y = neutraltweets, opacity = 0.95, name = "Neutra
```

Sentiment mapping program

```
# from the classified tweets, use location + latitude & longitude to map results
for line in reads:
    #print line
    lat = line[2]
    lon = line[3]
    pop_up = line[1]

    if line[0] ==  'Positive Tweet':
        folium.Marker([lat, lon], popup=pop_up, icon=folium.Icon(color='green'))
    elif line[0] == 'Negative Tweet':
        folium.Marker([lat, lon], popup=pop_up, icon=folium.Icon(color='red')).a
    #elif line[0] == 'Neutral Tweet':
        folium.Marker([lat, lon], popup=pop_up, icon=folium.Icon(color='yellow')
```