An N-Gram Determination of Twitter User Sentiments

C. Anyanwu[*], C. N. Udanor

Department of Computer Science, University of Nigeria Nsukka, Nigeria

Abstract: - **Today, due to the high reachability and popularity of social network sites, it is used by organizations for planning and mobilizing events for public demonstrations and protests. Social networking sites are sometimes used to launch harmful statements against individuals and groups. Some of its effects are an incitement to harmful practices. In order to overcome such activities on social network, there is a need to track and tackle hate speech on social networks and provide a way of apprehending hate speech mongers. The collection and documentation process of hate speech digital evidence online should be optimized efficiently. An automated system is needed to be able to implement such regulations. The aim of this paper is to determine what constitutes hate speech through a literature survey and identify their sources using an n-gram classifier, as well as proffer solution on monitoring and determination of hate speeches on Twitter, and perpetrators through the use of Twitter opinion mining. Results suggest that negative sentiments in form of hate speech is on consistent increase, given that the handles examined had between 20 to 70 percent negative sentiments contents.**

Keywords: **Twitter, Social networks, Sentiment analysis, Hate speech, Social network analytics.**

I. INTRODUCTION

Social media is an electronic communication platform that allows user constantly and continuously share information, make connections and freely express their thoughts across the world through different mediums at near zero marginal cost. On the other hand, social network is a social structure that connects people together with a common interest. Any user has right to post any kind of message and reach millions of users within a short period of time. These platforms has democratized message publication in that, any user can post message and interested user can obtain it. Today, due to the high reachability and popularity of social network sites, it is used by organizations for planning and mobilizing events for public demonstrations and protest [32]. Social networking sites are sometimes used to launch harmful statements against individuals and groups. Some of its effects are an incitement to self-harm practices [2]. The transformation of the social media has brought about many challenges that have manifested in a number of ways. Hate speech is an example of such challenges [3]. According to [4], there is no specific definition of hate speech on social media. But in broad terms, hate speech can be seen as any form of activity that has no meaning other than communication that expresses anger for a person or group of individuals, especially in situations in which the expression is likely to cause violence that is based on age, race, sexual harassment, political rights or gender. Anything tweeted can reach a huge number of people and its effect can be great extensively.

Hate speech is any communication that expresses anger or hatred for a person or group of persons based on a particular attribute of the persons being targeted. For example, in all Western European, all states are currently prohibiting various forms of racist, sexist, anti-religious, homophobic, or other intolerant speech [34]. The Penal Code in India also enforces prosecution against hate speech [35], in Nigeria [14], criminalizing hate speech is as a result of it consisting of incitement to an act of encouraging violence, discrimination, and hostility. It is used to silence unfavorable opinions and suppress debate. Hate speech is also characterized by the degree of attention given to the content and tone of the expression on social media such as fighting words (racial or gender), threats, defamation, which includes libel and slander, feelings or attitudes of hatred, obscenity, and so on [4]. Hate speech is also used by immoral agents for propaganda, based on hate speeches in a form that always becomes viral. Identifying the source of the message and holding the user that is responsible for the act might be difficult [5]. Physical violence springs up from both spoken and written words. Hate speech has actually done more harm than good.

A white supremacist Dylan Roof in 2015, was said to have killed nine African Americans at a Charleston, just after he was received by the bible study group at South Carolina Church. His action was based on the white supremacist propaganda he read and downloaded directly into his brain on the internet. He is currently under a death sentence [6].

Hate speech is considered as a broad umbrella for different sorts of abusive user-created words even though concerned authorities are regulating freedom of speeches on social sites. Social media has played a key role by providing esteemed channel for communication to people around the world. Though people misuse this right of freedom by indulging themselves in spreading rumours and undue criticisms. In order to overcome such activities on social network, there is a need to track and tackle hate speech on social networks and provide a way of apprehending hate speech mongers. The collection and documentation process of hate speech digital evidence online should be optimized efficiently. An

automated system is needed to be able to implement such regulations [7] [8].

The aim of this paper is to determine what constitutes hate speech through a literature survey and identify their sources. The rest of the paper is organized as follows: this section continues with; a brief history of the social network, an overview of the social network, its merits, and demerits. Section 2, discusses hate speech under the literature review, its sources and the harms it has caused. While section 3 proffers solution to the monitoring and determination of hate speeches on Twitter and pinpointing the user using text mining.

1.2 A brief historical background of the social network

Social networking started in 1971 with email, and in 1994 Geocities was created. The Classmate was developed in December 1995, being the earliest Internet social networking website. In October 1998 Open Diary was created, after 3 years Ryze was developed in 2001, and later in March 2002, Friendster was developed. A year later, in 2003 LinkedIn was developed, and in June same year hi5 was created. Also in August 2003, MySpace was developed. Orkut came on board in January 2004, and in February 2004 Mark Zuckerberg came up with Facebook. By March 2005 Yahoo! 360^0 was developed, and in July 2005, Bedo was developed. Jack Dorsey and others in 2009 were inspired by the popularity of text messaging and that prompted them to create Twitter and today, Twitter has over 50 million users as it allows more than 140 characters of tweets. Finally, in 2011 Google+ was developed [1].

1.3 An Overview of Social network

The social network is a kind of social structure that connects people with family and friends in a social circle to freely make new friends, express, share thoughts or interests. It has almost become part of our daily lives, it has established new communication behaviour and structure in societies. The social network is also an articulation of social connections achievable among individuals known as nodes (individuals within the network), which are links (relationships or interactions), edges, or ties that are connected by one or more types of specific inter-dependency, such as friendship, retweet or dislike etc.

The diagram in fig 1 shows a network structure. Two nodes are connected to each other if they interact regularly. Ngozi regularly interacts with Godwin but not with Chika. Therefore, there is no connection between Godwin and Chika but Ngozi and Godwin are connected.

Ties are connections among individuals used for sharing knowledge, experience, and information. These ties can be strong, weak or latent/absent based on the extent of interaction between the two nodes. Strong and weak ties are relevant in the interaction of social networking but plays a different role in community formation.

Strong tie (close friends) are connections within a circle of close friends and family. Weak tie (familiarity) are bridges that connect two different social networks. Absent/latent ties are expected existing connection but don't really exist [2].

1.4 Profitability of social network

Everything in life has its merits and demerits, social network inclusive. Some of its merits and demerits include:

i.   Social network helps to instantly connect you to millions of people all over the world. As Facebook is used to stay in touch with old school friends, Twitter to meet new people never known from cities or large area, and Google to hang out with relatives and professionals.

ii.   It helps in easy and instant communication. Instead of relying on a call to contact someone, with your phone or laptop communication can start immediately with anyone on social platforms.

iii. It helps in the discovery of real-time news and information. These days, no one waits for news on television or newspaper in the morning. One can choose what one wants by customizing one's news and information discovery on social media.

iv.   It provides enjoyment and general fun. Social network at times is fun. People turn on to it either at work or home. It is convenient and quite satisfying seeing what friends are up to, comments and likes show up on posts without asking them directly.

Demerits of social network

i.   There is no privacy. There is little or no control over what people say. That is why hate speeches find it a fertile ground to breed. And with so much sharing going on, sharing private data with the public can cause a lot of problems that cannot be solved at times.

ii.   Leads to distraction and disruption of sleep. Since social networking is performed mainly on mobile phones and laptops, people are distracted with news and messages they receive, thereby avoiding responsibilities. It has fed people with procrastination habit. It has also negatively affected one's ability to sleep well at night.

iii. It wastes time and money.

iv.   Leads to a major cause of hate speech. Since people are now connected all the time online, people use online interaction as a substitute for face-to-face interaction. This freedom of expression is responsible for the major changes in our society today. This transformation of social media has also led to abuse on the profitability of the social media, manifesting mostly in hate speech, cyberbully and predation of various sort. [33]

II. REVIEW OF RELATED LITERATURE

The term hate speech in this work is used since it is considered as a broad umbrella for different sorts of abusive user-created words that are addressed in individual works. Hate speech is one of the most frequently used expressions for this phenomenon which is even a legal term in some countries [9]. This is not surprising because hate speech has relatively been unchecked on social media sites for years, until recently when it has brought about many problems such as harassment, physical threats, and embarrassments among others. Social network and social media are tools that work together for social media strategy. These platforms have brought alongside many challenges. Hate crimes/speech are examples of such challenge among others. A report has shown that 60% of hate crimes and largest groups of social network users are youths [7]. A study from the Pew Research Center shows that 41% of American adults have experienced some form of harassment online [10]. The author in [4], argued that the harm in hate speech is mainly as a result of written speech rather than spoken speech. While [11], stated that hate speech lies with the freedom of expression, liberty, equality group right and dignity. They added that any objective definition in a computer program that can be easily implemented can be contested. Authors in[12], stated that a lot of hate promoters' group use well known social network websites to spread extremist content among their viewers. From the findings of [13], after examining issues on free speech that are posted by people on Facebook and attitudes towards delicate messages. He found out that hate message/information is on the increase and more people are joining.

Most countries are suffering from the problem of hate speech, which Nigeria is not an exemption. Hate speech has led to violence and loss of lives. According to Yemi Osinbajo, the Vice President of Nigeria who stated that hate speech is a terrorism spice. He proclaimed hate speech a crime at the National Security Conference in Aso Villa, Abuja. In an attempt to control hate speech, the federal government has drawn a line on hate speech [14].

Due to the relationship between social networks and hate speeches, an expert on the history of free speech and a professor of criminal justice [15], in a comprehensive summary of a social and hate speech controversy in the United States, primarily focused on the social context of prejudice, discrimination as a political issue and intergroup relations. It was centered on hate speech definition and various proposals in the efforts to control hate speech that have risen over the years.

As a result of various political, cultural and historical reasons, hate speech has been a long-standing problem in the world and has prompted much attention in recent times. The less advanced form of hate speech as argued by [11] has led to the growth of more sophisticated reading of symbols and signs which are hard to examine and remove parts that are considered offensive or dangerous. He demonstrated how these symbolic codes are historically and culturally imbued with meaning by using Nazi symbols appropriation and development as an example. He finally argued that hate can be expressed and experienced symbolically.

Hate and racism in most social networks are related. There has been several types of research/studies that have been investigating this site. The networks were analyzed systematically by racism group and found out that these sites have a decentralized structure. The software was used by the authors to automate the analysis of the content of hate mongers websites and their links on hate speech. They observed that the main objectives of these websites were to share ideas [16].

A review of online arguments and an overview of United State (U.S) currently for and against hate speech, internet regulations, and important jurisdiction was presented. In [17], an attempt to find solutions to limit hateful speeches on social media, they used YouTube as an example to explore the ability and type of the problem. They examined and evaluated the approaches that were proposed by other communication and legal scholars on whether or how to reduce online hate speech and sought solutions to reduce online hate speech. Their results tend to reduce hate speech on social networks site.

2.1 Conceptual Background

Sentiment analysis and hate speech are related closely, usually negative sentiment is connected to hate speech message. Because of this, the relatedness of hate speech and sentiment analysis are acknowledged by several approaches. **Sentiment** refers to the activity of five sense organs (eye, skin, ear, nose, and tongue) associating them with Feelings and emotions. That is the emotional effect of the text on the reader. Sentiments add content to the social network without it, text can be misleading. The author in [36] advised:

> *"Speak with caution. Even if someone forgives harsh words you've spoken, they may be too hurt to ever forget them. Don't leave a legacy of pain and regret of things you never should have said."*

The arms of sentiment are positive, negative or neutral depending on the words expressed within a given post on a particular subject. Understanding the words and classifying sentiments into its categories is necessary to be able to perform sentiment analysis. This is possible by using sentiment lexicons because different types of sentiment lexicons exist that have words that are classified to have sentiments that are either positive or negative. [14].

In positive words, positive sentiment is attached to it such as likes, comments, shares, kindness, love, happiness. Also, negative words have negative sentiment attached to them such

as hate, discrimination, sadness, hide and block a post. But when no emotion is applied, then they are neutral sentiment such as a click for checking out an image. Human Natural Language is a dedicated area in computer science that deals with the challenge of judging sentiments usually called Natural Language Processing. This area deals with the analysis of sentiment by extraction of information from some texts known as sentiment analysis. The sentiment is classified at various levels such as text level and sentence level (involves two steps: subjectivity classification and sentiment classification and feature level classification) [14].

Sentiment analysis is the analysis of opinion (such as like/dislike/mixed/ don't know), feelings (such as happy/sad/angry), satisfaction rating using Natural Language Processing (NLP), text analytics to determine, extract, analyze online social conversations and to determine the deep context as they apply to a topic. Sentiment analysis and opinion mining (a form of sentence, a feature whose work is to identify them as either negative, positive or neutral) are techniques that can be used to extract information based on your own opinions in text documents. To find the overall contextual polarity in any provided text, sentiment analysis is used. [37].

2.2.1 Social Network Analytics

It has become a necessity to find an effective and efficient way of identifying hate speeches, extracting text and learning about their relationship in order to examine the activities of hate speech. A set of methods has been provided on social network that can be used to analyze patterns of the structure of social entities. Hate speech is also increasing as a result of the growth rate of social media content. Recently, sentiment analysis has attracted lots of attention on Twitter because of its wide application on public and commercial sectors [18].

The fundamental concept of network analysis and a range of methods that are currently used was reviewed, as well as issues associated with the collection of data, network analysis, and individual level comparison and analysis was also discussed. In order to avoid common pitfalls in the application of network methods, suggestions were also made in [19].

Several types of social networks (SN) were enumerated by [20] as Simple SNs, temporal SNs, large scale SNs, virtual SNs, Multi-layered SNs, multidimensional SNs, heterogeneous SNs (with two or many different types of nodes) and much more. They also stated that some social networks have their own methods and measures developed. But the suitable analytical methods for the above SNs is SNA. Further stated, SNA may be distinguished in several real application domains which include: nodes classification for marketing purpose, organizational structural evaluation compared with communication structures in companies, hidden knowledge acquisition recommender systems for supporting users in web 2.0, web forum analysis of social groups and finally, their evolution prediction. They extracted

social communities and used appropriate algorithms and structural measures (SNA) for their statistics, calculation, and data mining methods in their work. [20]

Social Network Analysis is the measuring and drawing of various characteristics of distributed patterns of relational ties. It provides both mathematical and visual analysis of human relationships to represents the description of networks systematically. According to [38], SNA is a method for examining social structures and that is why it is referred to as structural analysis. Also, SNA is one of the methods used to identify, locate entity, examine and analyze the network dynamics of these structures which can be presented either as a multigraph or graph.

To examine and analyze the usefulness of information from huge data on the social network, special graph based mining tools will be required to easily model the structure of the social networks. A number of such analysis features, tools and benefits were listed by [22]. His work also presented a comparative network analysis tools such as network, Gephi (for interactive visualization and exploration of networks), Pajek (software for drawing networks, analytic capabilities, e.t.c), IGraph (for creating and manipulating graph, which is based on the platform, algorithms complexity, Graph types, execution time, features of the graph and the input file format).

The increase in hate speech on social network motivated [9] to present a survey on key approaches that can recognize these speeches automatically using Natural Language Processing and limits to those approaches. They focused on outlining existing approaches in a systematic manner (character-level, token –level approach) and extraction of features (such as simple surface features, word generation, sentiment analysis, lexical resource, knowledge-based features and much more). These tasks (sentiment analysis and lexical resource) were seen as a supervised learning problem and stated that the supervised learning approach should be applied in classification and utilization for hate speech detection. They also argued for a benchmark dataset for the detection of hate speech for better comparability of different features and methods [9].

In 2016, a methodology was proposed by [23]. The author classified messages containing hate or violent speeches in a very precise way after an attack in 2015 against Charlie Hebdo in Paris. An important public reaction was generated after the disruptive events in social networks, which led to the craving for the study of hate message and violent communication on Twitter. The authors carried out a qualitative analysis using data mining method to classify the type of speech. Finally, they identified that the disruptive event was as a result of communication that shows textual patterns and spatial-temporal that was identified clearly. These motivations made the authors to propose a

methodology that classified messages containing hate or violent speech in a precise way.

With SNA tools, [13] applied content based SNA approach in their work to find out people's interests in a mailing list network. From the entire communication network, they extracted an overlapping topic that is related to sub networks. Later, combined it with Text Mining (TM) to find out the extent to which sharing interest is connected with communication in two R-mailing lists (R-help and R-devel) where related questions of all kinds of applications and developments were discussed to describe people's interest. The authors found out that the communication efficiency is high in active mailing list than mid-active list. In order to find the relationship between communication and interest sharing, they recommended that only the subject should be used as more noise are contained in its content.

Research that explores the innovative and cross-methodological computational approaches developed an investigative study of criminal behavior [24]. They combined SNA and Agent-Based Modeling (ABM) into CrimeMiner in order to support the study of criminal organization. An experimental framework was developed that seamlessly integrates visualization and document enhancement [24].

Network analysis and web mining techniques have been widely used in the analysis of website contents and structures of hate groups on social media, these techniques have not been applied in blogs in the study of hate groups. To address the problem, a framework was proposed consisting of four modules: network analysis, information extraction, blog spider and visualization. The framework was applied to Xanga which is a popular blog hosting sites to identify and analyze 28 selected set of anti-blacks hate groups. In these groups, their analysis result shows some interesting topological characteristics, demographical and two large communities were identified on top of the smaller ones. They suggested that the framework proposed can be applied to blog analysis and can also be generally used in other domains. [25]

In 2014, a comprehensive review of software tools was provided [26] for social networking media such as newsgroups, Wikis, blogs, chat and much more. The review was written for researchers that seek to analyze scrapping and analytics on social media in their research. They included an introduction to the scraping, data cleaning, storage and sentiment analysis of social media. As a result of the availability of social networks APIs. Such as News Services APIs, Twitter, and Facebook, they provided a methodology and critique of social media tools.

2.2.2   Sentiment Analysis Techniques

Sentiment analysis is a knowledge discovery technique that is developed from data mining. Sentiment analysis and opinion mining are methods used to identify and extract subjective information in text documents. In [28], with sentiment analysis, the overall context polarity on any topic that is provided by its author can be found. Also, an opinion can be in any form of feature or document, sentence and the challenging work of opinion mining is sentiment classification that can be labeled as positive, negative or neutral which is done by guessing opinion about anything.

The purpose of sentiment analysis is to reveal the opinions of people on a specific topic. The researchers [31], applied sentiment analysis to identify the opinions of people on Twitter which they compared between sentiment analysis result, concepts and applications involving four strategies of social media used by government. Results show that sentiment analysis can contribute to the practices of social management.

Following the widespread and emotive event in Woolwich on the killing of Drummer Lee Rigby, United Kingdom led to a public reaction on social media which was interpreted as hate speech motivated [27] to use supervised machine learning to classify, train and test samples of tweets collected as hate speech which focused on ethnic or religion. Out of 2000 tweets, 75% was selected and others discarded. They applied Stanford lexical parser and context-free lexical model using features derived from dependencies in tweets of each content in the feature selection phase. To find the best features for hate speech classification, they combined rule-based, spatial based and probabilistic classifier and achieved 0.95 as an overall f- measure using features from content on tweet [27].

Despite the widespread of the problems posed by social media contents, reliable solutions are lacking for detecting hate speech. In an attempt to detect hate speech on social media, different methods were examined. Out of which, a supervised classification method was applied using a dataset that was recently released. Using word n-grams, character n-grams and word skip grams as features in the system. Their result shows that differentiating profanity from hate speech is a challenging work. [31]

Author [32], developed a reliable tool for the detection of hate tweets. Their approach was based on Naïve Bayes Classifier Algorithm that detected hateful speech/tweet. Naïve Bayes involves tweet acquisition and streaming to remove unwanted parts of speech using Tweepy API and pre-processing. N-grams was also used for documents structuring, and tweet classification to classify hateful speech that makes use of contents that are produced by self-identifying hateful communities from twitter and evaluation using Naïve Bayes. From their experiments, the result shows significantly that better performance was achieved using Naïve Bayes with precision, recall, and accuracy values of 58 %, 62%, and 67.47% respectively, than existing methods in hate speech detection algorithms.

The researcher [33], implemented hate speech by combining the above two approaches. (Supervised learning approach and Naïve Bayes in advanced method). Religious beliefs are considered as a target domain for the spread of hate speech.

To identify hate speeches, the researcher also devised a methodology to be able to filter tweets on Twitter. In performing their work, they investigated supervised learning algorithms utility such as Naive Bayes (NB), Support Vector Machine (SVM) and k Nearest Neighbors (kNN). The first categorized opinions on tweets by applying SVM, NB and kNN and then their sentiment polarity was found. Their result shows that in both cases, better performance is shown in SVM than kNN and NB [33].

Our approach is different from the above work because it is based on dictionary filtering in python flask. We use real-time tweets streaming text to determine hate speech and analyze the sentiment polarity.

III. METHOD FOR THE DETERMINATION OF HATE SENTIMENT

Our Python application uses the Twitter API to stream tweets from popular Nigerian Twitter handles in politics, ethnicity, religion, etc. Tweets were extracted from most frequent hashtags/words and the messages collected were divided into positive and negative classes to be matched with the N-gram dictionary. Unsupervised classification was applied to mining (positive or negative sentiments) in the unstructured texts. In quantitative terms, the application was able to identify the sentiments as either negative or positive by the computation of text to vector. Python dictionary (Dictionary N-gram Word Matching Method), a Python function that connects to open sourced Twitter API and SQLite 3 database was used to build our data collection server, and word cloud was used for the analysis. The N-gram dictionary was used in real-time to determine the polarity of sentiment and possible presence of hate sentiment in the streamed tweets. The tweets were streamed as unstructured texts and stored in the application's memory for immediate processing, ensuring there is no data redundancy. However, the user may want to store his result in a database. Python programming language was used for indexing due to its versatility and agility.

Tweets from Twitter were considered a valuable resource for determination of hate sentiments due to the following reasons:

i. Tweets are more descriptive, longer and publicly available, unlike Facebook.
ii. Twitter is meant for people you want to communicate with not for family and friends.
iii. The Twitter message contains a time zone that is often associated with the tweet.
iv. Twitter is a tool for the promotion of brands and products as the producers cannot push their products to a consumer without the consumers pushing back at them. For instant customer service, Twitter is the easiest and quickest way to contact companies.
v. Twitter offers an up to the minute analysis of an outbreak due to its frequent high message posting.

vi. Twitter has several sets of users and is also loved by celebrities.

3.1 Method Description

The approach involves the following steps:

i. Data collection using Twitter API: Large datasets of twitter is publically available. Thus, twitter data was extracted through the Twitter API.
ii. Data preprocessing: this involves streaming of data through python dictionary and word cloud.
iii. Tweets classification: the results from the above steps classifies tweets to either positive or negative.
iv. Graphical representation of sentiments: the sentiment analysis result is presented using pie charts.

3.2 Data collection using Twitter streaming API

Data collection is an essential requirement for training a classifier, but for our work there were no readily available dataset so we had to resort to real-time streaming of tweets using Twitter Application Programmer Interface (API). An API makes it easier for web services and computer programs to interact. APIs are provided by web services to be able to interact with their services. Four pieces of information are needed by a user application to establish a connection with the Twitter Streaming API: API key, API secret, Access token, and Access token secret. These parameters were first obtained by the researcher, using his Twitter handle before the proposed application could gain access to Twitter public data.

3.3 Tweets pre-processing

Using our application, tweets were collected from Twitter via the API based on the user query. The tweets collected were subjected to five phases; collection of data, data cleaning, data transformation, data visualization and selection of the algorithm. Preprocessing is done in the first phase. The raw text is a bit messy so before performing any analytics, the data needs to be cleaned. Cleaning the data requires removing some unwanted characters such as hashtags, https, stop words, special characters and replacing them with empty strings or none, as well as converting all texts to lower case.

3.4 Classification of tweets

Sentiment analysis uses text analytics in understanding the polarity of sentences. Tweets were extracted, and by applying part of speech (POS) tags to the tweets, nouns and verbs were extracted. The goal of the classification stage is to assign the polarity of a sentence in a document as negative or positive. To build our classifier, we used a library of python called Dictionary N-gram Word Matching Method (DNWM) to match the words retrieved from tweets by the system. Python is a powerful scripting language, which puts applications together, and integrating a system is quick and efficient with its support for multiple programming paradigms. The tweets collected from the Twitter API were applied to the dictionary

classifier. The result from the dictionary classifier then separated the tweets into (positive and negative) tweets represented as numeric data (i.e. vectorization) and the sentiments represented in pie chart format, while the specific word sentiments were displayed using the word cloud.

3.4.1 Dictionary N-gram word matching method

N-gram is a sequence of token that is derived by breaking sentences into uni-gram, bi-gram or tri-gram (words or phrases) that may appear in a bag of words. In python, an input sentence is just a string of characters that can be used to quickly generate n-grams. The Dictionary classifier is a unique pattern which is the main algorithm. We upgraded it using DNWM. The DNWM is a new dictionary method that is implemented in this research work, no built-in function that can handle this task. This is due to the fact that we want the keys to be ordered and the values to be unique, while dictionary keys are inherently unordered and the values are not checked between keys.

The following is the algorithm scripted for this work.

Dictionary={"Key1": "String", "Key2": "String2"} #Our test dict

print Dictionary

last_numeric = 2 #keep track of the highest numeric key for simplicity

strings = ["String2", "String3", "String4", "String"]

for x in strings: #Loop through the strings we want to match

if x not in Dictionary.values():   #If string not value in dict, add the string

Dictionary["Key {}".format(last_numeric+1)] = x   #Add string with key formatting

else:

print x

last_numeric += 1   #Keep track of highest numeric key

print Dictionary

"""Prints

{'Key2': 'String2', 'Key1': 'String'}

{'Key2': 'String2', 'Key1': 'String', 'Key 3': 'String3', 'Key 4': 'String4'}

Data dictionary was made with a list of words (positive and negative),

    Number of positive words: 2006

    Number of negative words: 4783

The proposed dictionary of words which we have used to differentiate tweets into positive and negative are listed in Table 1. Fig. 1 shows the implementation architecture of the system.

Table 1: The proposed word dictionary for the distribution of tweets (positive and negative)

|	Proposed Word Dictionary
1	*Positive*: abound, abundance, admire, honored, love, hopeful, rich, cool, *Negative*: abnormal, abuse, addict, ail, suck, bad, terrible, hate
2.	*Positive*: cute, crisp, deft, delight, delicacy, dote, divine, good, enjoy *Negative*: buggy, bull, cancer, careless, cold, choke, comical
Statistics based	*Positive*: pretty, love, best, wonderful, good *Negative*: hate, stupid, bad, waste
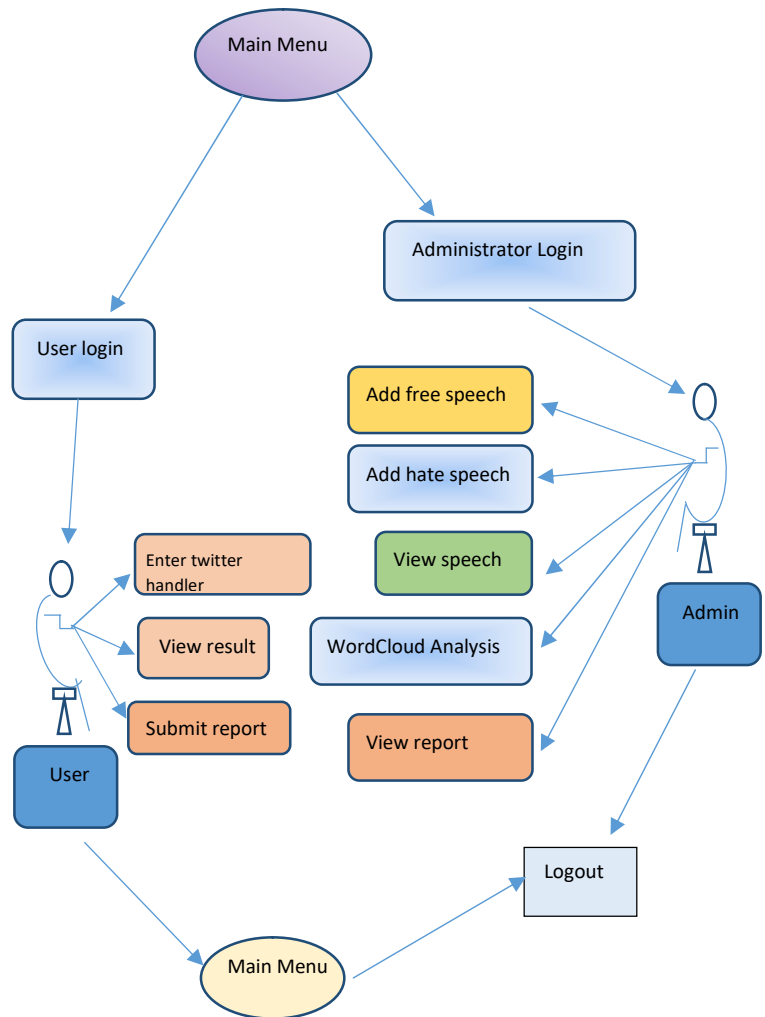
Fig. 1: implementation architecture

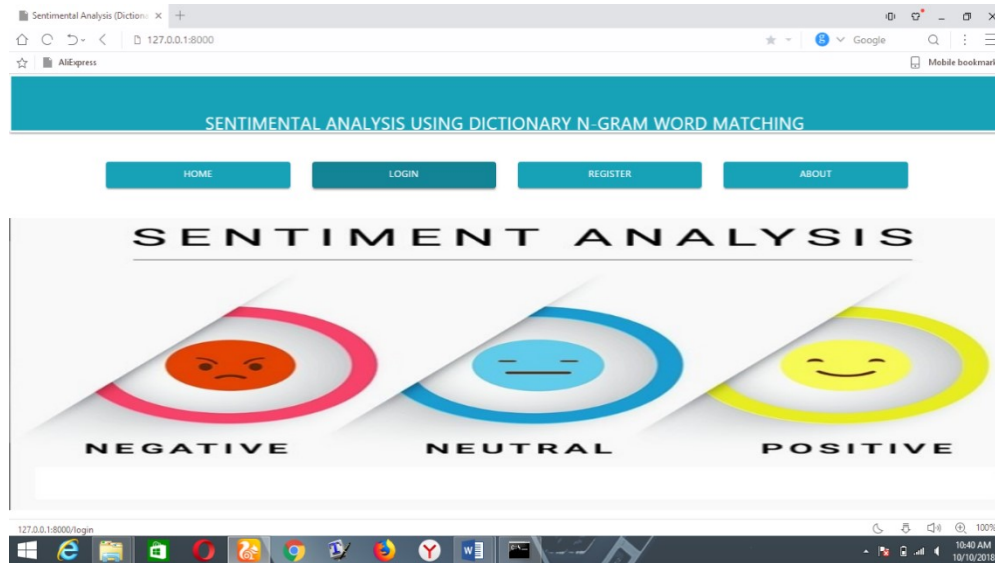Fig. 2 shows a screenshot of the landing page of the system.

Fig. 2: A screenshot of the landing page of the system

The software module is as follows:

i. Home page: the home page is designed and implemented using html language with bootstrap. On this page you have the login module, register and about.

   a. Login module: it has two input textbox and button. This is where every user present their parameters, handler, password and access is granted to the user to the dashboard

   b. Register: this is where user enter their full name, twitter handler and pin and access is granted to the dashboard.

   c. About: this talks about the topic.

ii. Dashboard: this is a page where all registered user access after proper login. On this page you have the upload data and logout.

   a. Upload: it has three input text (handler, hate and free speech) and one buttons (upload). The upload button is used to submit data to the database admin.

   b. Logout: this is used for ending session.

iii. Admin: this is a separate module on its own that controls the entire functionality of the system. It monitors the activities of the sentiment analyzers. You have the hate speech, free speech, view data, wordcloud analysis and logout.

   a. Hate speech: this is used to add additional hate word to the dictionary (text file)

   b. Free speech: this is also used to add additional free word to the dictionary (text file)

   c. View data: all user handler submitted is viewed on the view data.

   d. Wordcloud: this is a link through which hate speech and free speech words are analyzed.

IV. RESULTS AND DISCUSSION

The results of the system are presented. Firstly, the result of the system is shown, then that of the Word Cloud. Finally, tables for the word cloud sentiment analysis of the software are provided.

4.1Results and Discussion

In the classification of tweets for sentiment analysis, the classifier has the capability to determine when a sentence is either positive or negative. The sentiment analysis software is able to stream real time tweets from different twitter handles, display the results in a pie-chart form, converting the sentiment scores to positive and negative percentages, respectively, as well as analyze the steamed data individually using sentiment analysis (Word Cloud). A Word Cloud text file was opened through which the tweets fetched were filtered from and analyzed based on hate speech and free speech contents.

Table 2 shows 12 counts of tweets collected from Twitter handles cutting across politics, racism, ethnicity, etc. The tweets summarize the opinions posted by users as at the time the data was captured. The tweets were analyzed to find the sentiment polarity, and percentage of hate sentiment. A pie chart distribution of the result in Table 2 is shown in fig 3.

Table 2: Summary of results from top Nigerian trending handles

Twitter handle	Ranking	Tweets		Type of media	Hate speech	Free speech	Percentage of hate sentiment
@TheTrendOnline	1	6,920		Politics	31	15	67.4
@cchukudebelu	2	219k		Racism	30	29	50.8
@lemisegra	3	22.2k		Racism	21	28	42.9
@profosibanjo	4	11.4k		Racism	10	40	20.0
@ait_online	5	104k		TV station	30	13	69.8
@oluremisonaiya	6	8223		Ethics	20	45	30.8
@TJAdethink	7	138k		Ethics	21	42	31.7
@neyopumpn	8	29.9k		Politics	48	14	77.4
@NelsonHavi	9	18.6		Politics	26	20	56.5
@emerieUduchukwu	10	92.8k		Politics	27	24	52.9
@ebirimobinna	11	15.2k		Ethics	25	15	62.3
@SFHNigeria	12	4413		Ethics	44	15	74.6

The results show a significant presence of hate sentiment ranging from 20 to above 70 percent across the handles examined. This calls for concern, the amount of negative sentiment people release on social network sites.

The screen shot of the classification of tweets by the system is represented in fig 4, which contains the pie chart distribution of hate speech and free speech, respectively, in addition to the actual textual contents of the tweets.
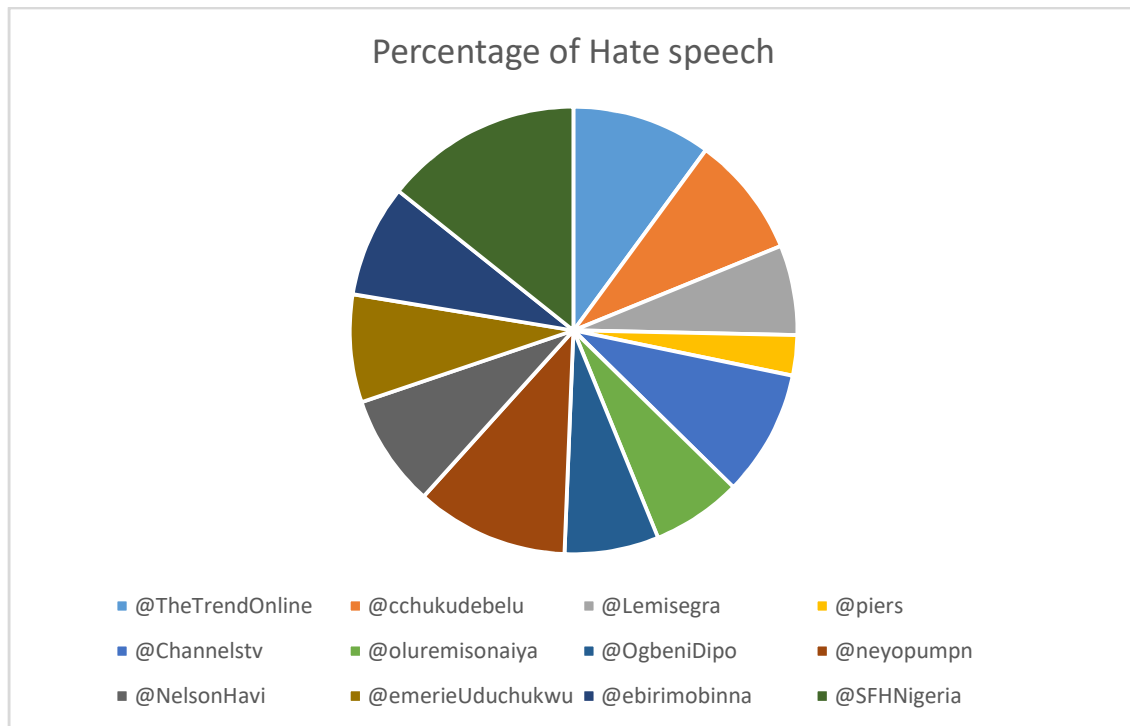

Fig 3: Results of count of tweets from different twitter handles

The pie chart in fig. 3 above represents the total number of words in the test data that has been classified as negative

sentiment from the classifier. The result shows that there is a consistent growth in hate contents on social media.

Fig 4: Screen shot of analyzed tweets on politics

The streamed data was also analyzed individually using sentiment analysis Word Cloud as shown in fig 5. Also, opened a WordCloud text file through which the tweets fetched are filtered and analyzed based on hate speech and free speech. The pie chart representation and WordCloud analysis are shown below.

Fig 5: Word Cloud analysis of the tweets

V. CONCLUSION

It is established in this study that negative sentiments which may manifest as hate speech is on the increase on social media. The study shows that peddlers of hate can actually be monitored with the contents of their messages stored as evidence. The findings can be used by government and law enforcement agencies to monitor user behaviours, pin down hate mongers and sanction them accordingly as the law may regulate.

REFERENCES

[1] Merriam Webster, "Dictionary and Thesaurus" (http://www.meriam_webstercom/dicionary/social%2520media), www.meriam_webster.com.

[2] S. John, C. Peter Carrington, "Social Network Analysis: An Introduction". http://dx.doi.org/10.4134/9781446294413.n2

[3] S. Leondro, M. Maniack, C. Denzil, B. Fabricro, W. Lngmar, (2016), "Analyzing the Targets of Hate in Online Social Media". Proceedings of the Tenth International AAAI Conference on Web and Social Media (ICWSM).

[4] W. Jeremy, "The Harm in Hate Speech", Harvard University Press, pp. 304. Reviewed by Brian Leiter, Univ. of Chicago. July 19, 2012.

[5] Wikipedia, "Hate Speech". https://en.wikipedia.or/w/index,php?title=hate_speech&oldjour_gradetds/15

[6] J. Murphy, "A Brief Analysis of the Free Speech vs. Hate Speech Debate_Stand", December 11, 2017. https://www.standleague.org/blog/a-brief-analysis-of-the-free-speech-vs-hate-speech-debate.html

[7] J. Levin, & J. McDevitt (1993), "Hate Crimes: The Rising Tide of Bigotry and bloodshed". New York: Plenum.

[8] S. Agarwal, and A. Sureka. A Focused Crawler for Mining Hate and Extremism Promoting Videos on YouTube. In Proceedings of the 25th ACM Conference on Hypertext and Social Media. Santiago, Chile, pp. 294-296, 2014.

[9] S. Anna, W. Michael, "A Survey on Hate Speech Detection using Natural Language Processing", Proceedings of the Fifth International Workshop on Natural Language Processing for Social Media, Valencia, Spain. Pages 1-10, April 3, 2017.

[10] M. Barthel, E. Shearer, J. Gottfried & A. Mitchell, "The Evolving Role of News on Twitter and Facebook". Pew Research Center's Journalism Project. Retrieved 2016-02-04.

[11] S. Walker, "Hate Speech: The History of an American Controversy". Lincoln: University of Nebraska. 1994.

[12] S. Agarwal, and A. Sureka. A Focused Crawler for Mining Hate and Extremism Promoting Videos on YouTube. In Proceedings of the 25th ACM Conference on Hypertext and Social Media. Santiago, Chile, pp. 294-296, 2014.

[13] B. Angela, F. Ingo, H. Kint& M. Patrick, "Content Based Social Network Analysis of Mailing Lists". R, Journal. Vol. 3/1, June 2011.

[14] The Nation Nigeria, "Hate Speech": In Editorial, Aug 30, 2017.

[15] R. Whillock& D. Slayden,"Hate Speech". Thousand Oaks, CA: Sage Publications, Inc. Introduction. pp. ix-xi.1995.

[16] Y. Zhou, E. Reid, J. Qin, H. Chen & G. Lai, "US Domestic Extremist Groups on the Web: Link and Content Analysis", IEEE Intelligent Systems. 20(5), 44-51. 2005.

[17] Ring, Caitlin Elizabeth, "Hate Speech in Social Media: An Exploration of the Problem and its Proposed Solutions." A Journalism & Mass Communication Graduate Thesis and Dissertations. 15. (https//scholar.colorado.edu. 2013.

[18] H. Saif, Y. He, M. Fernandez and H. Alain, "Contextual Semantics for Sentiment Analysis of Twitter", Elsevier B. V., 7th March, 2015.

[19] T. Carter Butts, "Social Network Analysis: A Methodological Introduction". Department of Sociology and Institute for Mathematical Behavioral Sciences, University of California, Irvine, Irvine, CA 92697-5100, USA. 2008.

[20] Y. Zhou, E. Reid, J. Qin, H. Chen & G. Lai, "US Domestic Extremist Groups on the Web: Link and Content Analysis", IEEE Intelligent Systems. 20(5), 44-51. 2005.

[21] Ring, Caitlin Elizabeth, "Hate Speech in Social Media: An Exploration of the Problem and its Proposed Solutions." A Journalism & Mass Communication Graduate Thesis and Dissertations. 15. (https//scholar.colorado.edu. 2013.

[22] N. Akhtar, "Network Analysis Tools", Fourth International Conference on Communication Systems and Network Technologies, DOI 10.1109/CSNT.2014.83, 2015.

[23] Fernando Miro-LIinares et al, "Cyber Hate Speech on Twitter: Analyzing Disruptive Events from Social Media to Build a Violent Communication and Hate Speech Taxonomy". Int. Journal of Design & Nature and Ecodynamies. Vol. 11, No. 3. Doi: 10.2495/DNE-V11-N3-406-415, July 2016

[24] N. Lettieri, A. Altamura, D. Malandrino& V. Punzo, "Agents Shaping Networks Shaping Agents: Integrating Social Network Analysis and Agent Based Modelling in Computational Crime Research". Springer International publishing EPIA 2017. LNAI 10423, pp. 15-27, Doi.10.1007/97 -3-319-65340-2-2. 2017.

[25] M. Chau, X. Jennifer, "A Framework for Locating and Analyzing Hate Groups in Blogs",

[26] B. Batrinca, P. Treleaven C., "Social Media Analytics: A Survey of Techniques, Tools and Platforms", AI &Soc, 30:89–116, DOI 10.1007/s00146-014-0549-4. 2015.

[27] P. Burnap, M.L. Williams, "Hate Speech, Machine Classification and Statistical Modeling of Information Flows on Twitter: interpretation and Communication for Policy Decision Making". In Proceedings of the Internet, Policy & Politics Conferences, Oxford, England, 2014.

[28] B. Liu, L. Zhang, "A Survey of Opinions Mining and Sentiment Analysis in Mining Text Data", pp. 415-463, 2012.

[29] P. Fornacciari, M. Monica and T. Michele, "Social Network and Sentiment Analysis on Twitter: Towards a Combined Approach", KDWeb.2015

[30] Merriam Webster, "Dictionary and Thesaurus" (http://www.meriam_webstercom/dicionary/social%2520media), www.meriam_webster.com.

[31] D. Oliveira Jose Silva; B. Paulo Henrique de Souza; P. Jose Roberto and B. DanielyAparecida, "The Application of the Sentiment Analysis technique in Social Media as a Tool for Social Management Practices at the Government Level", Rev. Adm. Publica[online]. 2019. Vol. 53, n.1,pp. 235-251. ISSN 0034-7612. http://dx.doi.org/10.1590/0034-7612174204

[32] S. Muthiah, B. Huang, J. Arredondo, D. Mares, L. Getoor, G. Katz N. Ramakrishnan, (2015), "Planned Protest Modelling in News and Social Media". In Proceedings of 29th AAAI Conference on Artificial Intelligence. Austin, USA, pp. 3920-3927.

[33] https://www.computerhope.com/jargon/s/socinetw.htm

[34] A G Noorani. Hate Speech and Free Speech. Economic and Political Weekly, Vol. 27, No. 46, p. 2456. Nov. 14, 1992

[35] C. N. Udanor, S.O. Aneke and B.O. Ogbuokiri, "Determining Social Media Influences of the Politics of developing Countries using Social Network Analytics". EmeraldInsight Program: electronic library and information systems, vol. 50(4): pp. 481-507. 2016

[36] Germany Kent, "Hate Quotes". Available at: https://www.goodreads.com/quotes/tag/hate-speech. (Viewed March 11, 2018).

[37] T. Zia, M. ShehbazAkram, M. Saqib Nawaz, B. Shahzad, Abdullatif M. Abdullatif, RazaUl Mustapha, M. IkramullahLali, "Identification of Hatred Speeches on Twitter" International Journal of Advances in Electronics and Computer Science, ISSN: 2393-2835, volume 4, Issue -1, Jan 2017.

[38] Mirigxin Zhang, "Social Network Analysis: History, Concepts and Research". In: Furht B. (eds) Handbook of Social Network technologies & Applications. Springer, Boston, MA. 2010.