**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**MONITORING SUSPICIOUS BEHAVIOUR ON SOCIAL MEDIA USING DATA MINING**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

****

**Supervisor:**

**Mr.**

**Akim Munthali**

**Author:**

**Terence Tachiona M147740**

**A RESEARCH PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF BACHELOR OF SCIENCE HONORS DEGREE IN COMPUTER SCIENCE**

**in the**

**Department of Mathematics and Computer Science**

**Gary Magadzire School of Agriculture and Natural Sciences**

**November 22, 2017**

**Abstract**

National security concern is the primary goal of any nation. Numerous security risks that exist in these networks include privacy violations, identity theft, and sexual harassment among others. Beyond being facilitators of human interactions, social networks have become an interesting target of research, providing rich information for studying and modelling user’s behaviour. Identification of personality-related indicators encrypted in Facebook and Twitter profiles and activities are of special concern in our current research efforts. This research project explores the detection of behavioural activity seen as suspicious on online social networks such as Facebook and Twitter. The research focuses on data mining algorithms as a way of monitoring the social media. The encouraging results of the study, exploring the suitability and performance of several data mining techniques, will also be presented.

**Keywords:** *Social Media, Data Mining, Machine Learning, Suspicious, Natural Language Processing, Text mining.*

**Declaration**

*Student*

I, **Terence Tachiona** Registration number **M147740** do hereby declare that the work contained in this research project is the result of my own investigation and research, except to the extent indicated in the Acknowledgements, References and by comments included in the body of the report, and that it has not been submitted in part or in full for any other degree to any other university.

**Student Signature \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Date\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

*Supervisor*

I, **Akim Munthali** confirm that the work reported in this dissertation was carried out by the candidate under my supervision as supervisor. This research project has been submitted for marking with my approval as supervisor.

**Supervisor Signature\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Date\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Dedication**

To my loving wife Benny, my precious parents Mr & Mrs Guvhu, and finally my siblings and in-laws.

**Acknowledgements**

This is the most difficult part of this thesis because so many people assisted me, but I must mention a few. I first give Almighty God the glory, for His mercy endureth for ever, and thank Him for keeping me alive until this day.

I express my profound appreciation to my supervisor, Mr Akim Munthali of the Mathematics and Computer Science Department, at Great Zimbabwe University. Thank you for your guidance, care, and constructive ideas throughout this research.

I thank my lovely wife, Benhildah for her moral and social support, patience and understanding during the many hours I disappear and become antisocial. I would like to thank Nyashadzaishe Muringani of MarketForce, who took his time and energy providing necessary technical and academic advice and all my friends, for their indefatigable advice.

I am grateful to the developers and maintenance staff of Anaconda, Facebook Inc., Twitter Inc. and Great Zimbabwe University towards the partial completion of this research is hereby acknowledged. However, opinions expressed, and conclusions arrived at, are those of the author and are not to be attributed to those mentioned above.

***That which you are seeking is seeking you.***

**Acronyms**

NLP – Natural Language Processing

ICT - Information and Communication Technology

NLTK – Natural Language Toolkit

ML – Machine Learning

API – Application Programming Interface

AI – Artificial Intelligence

FQL – Facebook Query Language

**CHAPTER 1: Introduction**

**Introduction**

Nowadays people are dependent on utilizing the web as it has turned into an extremely effective communication channel for people to share their knowledge, express their opinion, promote their products, or even educate each other, by publishing posts on online forums in form of text, audio, video and images. A great survey (Gladwell and Shirky 2011) shows that the popular social networking websites Facebook and Twitter have over 2 billion and over 800 million active users respectively. In 2010, according to the report by CNN, 75% of the news was forwarded to other people through email and 37% of the news items were shared on Facebook and Twitter.

Obar (2015) defined social media as computer-mediated tools for creating, sharing and exchanging of information, photos, videos and career interests in the virtual communities and networks. Social Media services have been defined as Web 2.0 applications and web forums that emphasize on online generated content and interoperability (Darcy, 1999). This has led to web forums, blogs, wiki and media-sharing websites that grant online users the capability to communicate and respond to opinions on daily news, political debates, tribal and religion discussions. Sutton (2009) formulates social media as diverse and developing online communication that allows the production and sharing of information. Crisis communication, public relations, radicalization, crime and illegal behaviour has been acknowledged as the roles by which social media is aiding online communication. Social media forums such as Facebook and Twitter have played a daily constant role in issuing daily political news, traffic updates, emergency warnings and alerts, supporting recovering efforts, situational awareness and real-time information.

**Table 1:** *Internet Traffic Report by ALEXA Website on 15 September 2017*

|  |  |  |  |
| --- | --- | --- | --- |
| Rank | Website | Rank | Website |
| 1  2  3  4  5  6  7 | Google.com  Youtube.com  **Facebook.com**  Baidu.com  Wikipedia.org | 8  9  10  11  12  13  14 | Amazon.com  Tmall.com  **Twitter.com**  Google.co.jp |

Research reveals that Twitter and Facebook are in constant use by activist groups for dispersion of their thoughts, viewpoints and beliefs. The Law Enforcement and Communication authorities need to setup the war on the vice so as to avoid future problems. Social Media and Crime has shown a clear relationship that someone knew of the possible crime or planning of the crime. Wilson et. al (2012) argued that “Facebook and other social media platforms need to be positioned firmly within the criminologist’s gaze not only because of the wealth of data these platforms provide but also due to their significant popularity across all age groups and their influence on how people interact and communicate with one another. This is because social media is now intimately interacting with many of the issues that we as criminologists can and should be engaging with.

Mining helpful data from those posts is critical for individuals to reveal the shrouded information. Content information mining calculations are utilized to identify criminal exercises and illicit postings. The system screens and examines online plain content sources, for example, web news, web journals, and so on for security purposes. This is through data mining algorithms. Data is regularly determined through the formulating of examples and patterns. These posts and comments are analysed for provoking posts by comparing the words in the posts with the set of sensitive key-words in the algorithm. Further the set of sensitive keywords are divided into different categories which include political, sexuality, religious, historical, racial discrimination.

**Fig 1:** *Basic Process of Suspicious Activity Detection*

**Research background**

The social media is an immense virtual space where to express and offer individual sentiments, impacting any part of life. The exponential headway in data and correspondence innovation has cultivated the making of web-based social networking, for example, Facebook, Twitter, LinkedIn, Instagram and Pinterest for much online discourse.

Unfortunately, malicious people take advantage of this technological achievement in the sense that they use it for illegal purposes. In social media, the users produce several and various formats of suspicious posts (text, image, video) and exchange them online with other people. The data in most social media sites are stored in text format, so in this work we will focus only on text posts.

Escalating crimes on social media alarms the law enforcement bodies to keep a gaze on online activities which involve massive amount of data. This will raise a need to detect suspicious activities on online available social media data by optimizing investigations using data mining tools

Social Media, web forums and online conversations and debates have brought to the fore illegal behaviour. Some of the illegal behaviour is due to the ethnic tensions which have polarized the country with each passing political season. Illegal behaviour was in the limelight during the 2008 post-election results, as more citizens were using phone short messages to fuel, share destructive propaganda and support politically based violence of the different political communities depending on where you were aligned. Current studies have shown the movement of illegal behaviour from traditional person to person phone message in to social media platforms such as Facebook and Twitter. There was a swift change of political bickering and political tension from the physical to the digital cyber space. This was fuelled much by the almost instant delivery of propaganda, agenda and tribal ethnic through social media forums such as Twitter and Facebook. Currently in Zimbabwe there is maximization of social media for campaign purposes, voter communication and civic engagement. This in turn saw the resurgence of online illegal behaviour which led into violent clashes between pressure groups such as Tajamuka, ThisFlag and law enforcement agencies such as Zimbabwe Republic Police. This resulted in the government enacting laws to control illegal behaviour both online and offline through the Cybercrime and Cyber security Act of 2017.

On social media, the clients deliver a few and different configurations of suspicious posts and trade them with other individuals. In such manner, the researcher saw it basic to build up a framework that screens the posts on these web-based social networking for conceivable illicit exercises that are in content configurations can be additionally utilized as proof for examination.

**Problem Statement**

In the age of information and internet, online data is growing in an exponential manner and human behaviour and social interactions are greatly influenced by the digital and multimedia technology. They exploit this communication channel for promoting radicalization, recruiting members and creating online virtual communities sharing a common agenda. These malicious users who take advantages of this technological feat to publish illegal and suspicious contents (images, videos, texts …) in order to exchange data online and share ideas that could affect the security of countries or institutions.

Because of the wrongdoing related multifaceted nature connections, the conventional techniques for observing and identifying unlawful exercises are outdated and expend additional time and human resources. Also, these strategies are not ready to acquire every single persuasive parameter in light of their high measure of human obstruction, in this way, the law requirement offices are searching for an insightful and deliberate approach for distinguishing suspicious conduct on social media.

Advent of these interacting networks, led to the increase of countless crimes, as they offer ease to criminal conversations and transfer of data (suspicious messages).As an example, the media sharing websites for YouTube allow to publish videos in relation to “how to create a bomb”. The social network Facebook and the micro blog Twitter also help criminals to coordinate and manage online suspect actions.

In this regard, the researcher built up a system which handled this issue using a data mining algorithm to identify suspicious exercises and illicit postings on Facebook and Twitter. The framework utilized a text data mining strategy to lessen numerous illicit exercises which are hung via social media.

**Objectives of the Study**

**Overall Objective**

The objective of this research is to develop a social media monitoring tool that detects suspicious behaviour on Facebook and Twitter through data mining on posts by users.

**Specific Objectives**

* To identify and analyze techniques used in illegal behaviour monitoring and select the best suitable.
* To demonstrate and test the application while providing analysis on the illegal behaviour on Facebook and Twitter in Zimbabwe.
* Scans user posts and tweets, views, comments, likes for criminal activity.
* Detects suspicious behaviour and automatically flag the content
* Allows admin users to read the flagged content details
* Determines if further steps are required

**Hypothesis**

Initially, the researcher will conduct an analysis of the existing and proposed systems. This preliminary stage is useful to the research as it determines the feasibility and plausibility.

The researcher will undertake requirements engineering that will enable system design. The design involves identifying and describing the fundamental system abstractions and their relationships.

In order to accomplish the design and implementation stage, the researcher is going to use the following tools:

* Python 3.6
* Jupyter Notebook
* Natural Language Tookit
* Facebook API
* Twitter API

**Justification**

Cyber criminals have been known to move from the cyber space to the actual physical world to promote, fund and finance violent crimes. Hence, the need for this monitoring tool to capture the digital evidence. The research proposal is a much needed approach to address the gap cyber criminals are taking advantage in the cyberspace arena. It’s critical and a much needed approach to check the cyber space as such acts of illegal behaviour are affecting and influencing different people, races, tribes and the entire country. The tool will assist law enforcement agencies to easily and readily make use of the features to capture data and digital evidence.

With this tool in hand law enforcement and communication authorities will be on high alert and be able to bring down offensive social media web forums and thus reduce potential political, ethnical and tribal conflict. After all, prevention is better than cure when it comes to civil war and internal conflicts.

This research study is imperative in various approaches to the researcher, Great Zimbabwe University and the world on the loose. To the researcher, this incorporates the change of the information on the advancement of frameworks that utilize information mining systems. The study is additionally basic to the researcher as it is as per the prerequisites of the degree program towards which the researcher is considering.

The study is also critical in honing of the researcher's expressive, research and developmental aptitudes which are of focal essentialness in the science, technology engineering and mathematics world.

To the university, the research will provide literature for review in future by other students and academic staff who may wish to undertake a research on a similar topic thus it can be used as a pedagogical tool for teaching artificial intelligence. Social media has been used as a communication channel in several circumstances, namely the Egyptian revolution, Boston’s attack, etc.

**Assumptions and Limitations**

The following assumptions and limitations are considered

* It is assumed that the cyber forensic investigator has access to the social media web forum or account or is friend/follows the person of interest. The web forums are also thought to be open to the public and easily accessible on the Internet.
* The potential risk of either a website or social media account being hacked is always present. With this in mind, it’s assumed that the social media web forums comments on a person of interest account or website are his/her publications and there was no hacking onto the user’s web forum or account.
* The researcher will have enough time and assets to finish the research project in a timely manner.

**Scope**

This investigation focuses on the improvement of a framework that downloads postings from Facebook ceaselessly and utilizes data mining algorithms to identify hotly debated issues and bunch writers into various gatherings utilizing word-based client profiles. The framework makes an investigation of text in Facebook and Twitter posts utilizing text data mining algorithms. The researcher confined the study to the overall aim and individual objectives of the research.

**Summary**

To total up, the researcher has presented an application of data mining which will help to curb the effect of posting abusive articles or comments on social network that is likely to harm the sentiments of society knowingly or unknowingly. The chapter defined the problem by focusing on background to the study, statement of the problems, research objectives and justification of the study, assumptions, scope and limitations.

The problem statement was presented in its local context and the overall objectives of the study have been stated. The chapter provided the rationale to conduct the study. The core of the study which involves mainly of a system capable of monitoring posts on Facebook and Twitter using text data mining techniques is relied upon to hold up under productive results.

**CHAPTER 2: Literature Review**

**Introduction**

The continual pursuit of offenders in the cyber realm is an ever evolving practice and cyber forensics had been in the forefront. The need for newly released and better forensic tool is a constant desire any cyber forensics expert. Planning, acquisition and reporting of cybercrime is more stretched when the digital evidence is not available locally on a hard disk drive or digital media, as the cyber forensics expert needs to rely on remote acquisition of illegal behaviour digital evidence from social media web forums that are not hosted locally. Cyber criminals are always finding need ways to propagate their agenda with the use of more sophisticated IT tools and applications. Thus, cyber forensic experts have to come with more advanced tools, applications and systems that allow them to catch the illegal behaviour mongers.

Rapport (2010) describes forensics for social media as any listening methods or solutions that allows technology and provides an algorithm for helping researchers and organizations to listen, collect data, document, analyse, interpret and respond to online conversations. This further collaborated by Branthwaite and Patterson (2011) who said “when compared to the traditional research approaches, the similarities that social media monitoring shares with quantitative research include large samples, numeric data and difficulty in assessing meanings, while among those it shares with qualitative approaches are the gathering of spontaneous views and opinions, and a need for rigorous semantic analyses”. Furthermore, social media monitoring can be thought to be a continuous daily process that offers insights, information, intelligence and communication for planners and research as found out by Hipperson (2010). Traditionally, social media monitoring and quantitative approaches consist of sampling and standardization in big data research.

Social Media data constitutes of what is known as big data. Branthwaite and Patterson (2011) argue that the research differences with qualitative approaches in big data samples is lack of direct contact, huge research targets due to huge data sets, differences in non-verbal cues, feedback, data collection and contextual information Zailskaite-Jakste and Kuvykaite (2012) demonstrated social media monitoring as function of gauging, listening and analyzing the online environment. They went ahead to formulate that it may serve the purpose of evaluation, tracking the success of the illegal behaviour message sent. Monitoring and evaluation of such web forums is paramount and critical for the general public, authorities and communication agencies, it through such initiative that illegal behaviour can be mitigated if not eliminated.

Social Media monitoring can be considered as active methods of engaging citizens in online interaction, gathering online information and analyzing the data for useful purposes. A case in point is the example where health officials engaged online with a sample population to collect, gather and analyze online data through a survey on an e-health website (Laakso, Armstrong, & Usher, 2012). The research study undertaken for this project included social media monitoring of organic and interactive online conversations and communications on web forums

**Reasons for Monitoring Social Media**

Monitoring is conducted to collect online data, analyze the data and make inferences on the analyzed data. It is from such research discussions that the general public, authorities and communication and regulatory authorities can be able to understand the online virtual environment that they are in (Zailskaite-Jakste & Kuvykaite, 2012). By understanding the environment it allows the general public to be educated and made aware of the dos and don’ts when it comes to illegal behaviour communication. The authorities, communication and regulatory bodies will be able to come up with laws, rules and regulations for reporting and capturing the illegal behaviour mongers. It allows a proper outline and structure for illegal behaviour monitoring and reporting. Kavanaugh et al. (2012) demonstrated the goal for social media monitoring is to allow a bird’s eye view and the phenomena of big picture evaluation and analysis. This is further collaborated by Bengston et. al (2009) when they conclude that monitoring allows a peep in the societal debate window and illuminates the stakeholder perceptions, wants and attitudes. The driving force for monitoring may be related to public safety (Kavanaugh et al., 2012), product analysis (Deluca et al., 2012), political opinions, public reactions to policies (Sobkowicz et al., 2012), identification of radical opinions (Yang, Kiang, Ku, Chiu, &Li, 2011), illegal behaviour, profiling (Keelan, Pavri, Balakrishnan, & Wilson, 2010) and falsified information (Campbell, Pitt, Parent, & Berthon, 2011).

Social networks are typically rich in text, because of a wide variety of methods by which users can contribute text content to the network. For example, typical social networks such as Facebook allows the creation of text content such as, wall posts, comments and links to blog and web pages. Emails between different users can also be expressed as social networks, which can be mined for a variety of applications. For example, the well known Enron email database is often used in order to mine interesting connections between the different characters in the underlying database. Using interesting linkages within email and news group databases in addition to the content often leads to qualitatively more effective results. Social networks are rich in text, and therefore it is useful to design text mining tools for a wide variety of applications.

While a variety of search and mining algorithms have been developed in the literature for text applications, social networks provide a special challenge, because the linkage structure provides guidance for mining in a variety of applications.

**Data Mining in a Nutshell**

One definition of data mining is identifying novel and actionable patterns in data. Data mining is also known as Knowledge Discovery from Data (KDD) or Knowledge Discovery in Databases, also abbreviated as KDD. Data mining is related to machine learning, information retrieval, statistics, databases, and even data visualization. One formal definition for data mining is found in Princeton University’s WordNet3 where data mining is defined as:

*“data processing using sophisticated data search capabilities and statistical algorithms to discover patterns and correlations in large pre-existing databases; away to discover new meaning in data”*

**Data Mining in Social Media**

The key idea behind data mining is finding new information in a data set that is hidden or latent. Data mining can help people better understand large sets of data. Supervised and unsupervised algorithms are used to identify the hidden patterns in data. Supervised approaches depend on some a-priori knowledge of the data (e.g. class labels). Unsupervised algorithms are used to characterize data without any prior instruction as to what kinds of patterns will be discovered by the algorithm. The variety of work accomplished to date pertaining to data mining of online social media data is accomplished with some version of either supervised or unsupervised learning algorithms. Determining whether a supervised or an unsupervised approach would be best depends on the data set and the particular question being investigated. Datasets can be generalized into three types: data with labels, data without labels, and data with only a small portion of labels.

Classification is a common supervised approach and is appropriate when the data set has labels or a small portion of the data has labels. Classification algorithms begin with a set of training data which includes class labels for each data element. The algorithm learns from the training data and builds a model that will automatically categorize new data elements into one of the distinct classes provided with the training data. Unlike classification algorithms, clustering algorithms do not depend on labelled training data to develop a model. Instead, clustering algorithms determine which elements in the data set are similar to each other based on the similarity of the data elements.

**Social media data**

Clearly, there is a large and increasing number of (commercial) services providing access to social networking media (e.g., Twitter, Facebook and Wikipedia) and news services (e.g., Thomson Reuters Machine Readable News). Equivalent major academic services are scarce. We start by discussing types of data and formats produced by these services.

**Types of data**

Although the researcher focused on social media, as discussed, researchers are continually obtaining new and innovative sources of data to bring together and analyze. So when considering textual data analysis, we should consider multiple sources (e.g., social networking media, RSS feeds, blogs and news) supplemented by numeric (ﬁnancial) data, telecoms data, geospatial data and potentially speech and video data. Using multiple data sources is certainly the future of analytics. Broadly, data subdivides into:

**Historic data sets** - previously accumulated and stored social/news, ﬁnancial and economic data.

**Real-time feeds** - live data feeds from streamed social media, news services, ﬁnancial exchanges, telecoms services, GPS devices and speech.

*Data also subdivides into:*

**Raw data** - unprocessed computer data straight from source that may contain errors or may be unanalyzed.

**Cleaned data** - correction or removal of erroneous (dirty) data caused by disparities, keying mistakes, missing bits, outliers, etc. • Value-added data—data that has been cleaned, analyzed, tagged and augmented with knowledge.

The four most common formats used to markup text are: HTML, XML, JSON and CSV.

* **HTML -** HyperText Markup Language (HTML) as well-known is the markup language for web pages and other information that can be viewed in a web browser. HTML consists of HTML elements, which include tags enclosed in angle brackets (e.g., \div), within the content of the web page.
* **XML -** Extensible Markup Language (XML) is the markup language for structuring textual data using \tag[…\\tag[to deﬁne elements].
* **JSON -** JavaScript Object Notation (JSON) is a text based open standard designed for human-readable data interchange and is derived from JavaScript.
* **CSV -** a comma-separated values (CSV) ﬁle contains the values in a table as a series of ASCII text lines organized such that each column value is separated by a comma from the next column’s value and each row starts a new line.

**Motivation for Twitter Data**

Twitter is a gold mine of data. Unlike other social platforms, almost every user’s tweets are completely public and pullable. This is a huge plus if you’re trying to get a large amount of data to run analytics on. Twitter data is also pretty specific. Twitter’s API allows you to do complex queries like pulling every tweet about a certain topic within the last twenty minutes, or pull a certain user’s non-retweeted tweets.

A simple application of this could be analyzing how your company is received in the general public. You could collect the last 2,000 tweets that mention your company (or any term you like), and run a sentiment analysis algorithm over it. We can also target users that specifically live in a certain location, which is known as spatial data. Another application of this could be to map the areas on the globe where your company has been mentioned the most.

Twitter data can be a large door into the insights of the general public, and how they receive a topic. That, combined with the openness and the generous rate limiting of Twitter’s API, can produce powerful results. To connect to Twitter’s API, we will be using a Python library called Tweepy, which we’ll install in a bit.

**Challenges in Social Media Data**

Presently, the individuals are fortunate enough to share and articulate their opinions and viewpoints concerning various aspects of life on a single message board, Social website data. This data actually signify massive virtual space, where anyone can hold discussion in the form of posted messages. User preferences are generally captured by analyzing their attitudes and behaviors mentioned on the websites as comments. To measure a user’s loyalty and to keep track of their sentiments towards any topic is achieved by monitoring the suspicious activities and discussions done through their posts on the social media websites. The main hurdle faced by researchers in doing so, is the lack of information retrieval and data analysis tools for real time data. The resultant database is quite huge and thus to extract desired knowledge from the large search space of online social media data, an intelligent data mining algorithm is required.

Moreover, the involvement of massive numbers of parameters in the search space makes the extensive search infeasible. Consequently, proficient search approaches are of imperative significance. Thus, the study and analyses of data from online social websites’ textual data consists of numerous issues and challenges. The data available is not in ready-to-use format. Some of the major problems are discussed below:

* **Grammar and Spellings**

Most of the users make a lot of semantic or even spelling mistakes when they post something on the web. Using datasets, these mistakes are processed during the pre-processing phase of any application.

* **Trustworthiness**

The number of user’s views on different subjects signifies the importance of the data on the web. Unfortunately, numerous fake accounts are made to give these posts a fake view and also fake reviews are also given to either push or to pull a post or an entity on the web platform.

* **Format**

Every other online website exhibits their own way or format for data posting and also the different users have different style of posting. For example: hashtag (#) is used when subjects are to be tagged or @ is used to refer different users. Hence, each website needs to be studied separately.

* **Language**

The option to post views or data in different languages is also available in online websites. Also the translator option is available to understand the other language’s post.

**Categories of Techniques Employed**

The online suspicious detection activities can be categorized under different heads depending on the way they handle the data. Researchers have tried to group the developed techniques for monitoring suspicious discussions based on different criteria’s. However, the one proposed by Murugesan, Devi et al. 20162, received the acceptance. These categories are presented below with their specifications and related work.

* **Brute Force Algorithms** In brute force strategy type, the relations between inflected and root forms are contained by the stemmer’s lookup table. A word is stemmed by querying the table when inflection matching to stem is found. When matching inflection is found, the root form associated with it is returned.
* **Matching Algorithms** Matching algorithms use stem database (Example: a document set containing stem words). The algorithm searches the stem database for a match of the word that needs to be stemmed.

**Social media monitoring tools**

Social media monitoring tools are sentiment analysis tools for tracking and measuring what people are saying typically about a company or its products, or any topic across the web’s social media landscape. In the area of social media monitoring examples include: Social Mention, (http://socialmention.com/), which provides social media alerts similarly to Google Alerts, Lithium Social Media Monitoring; and Trackur, which is an online reputation monitoring tool that tracks what is being said on the Internet. Google also provides a few useful free tools. Google Trends shows how often a particular search-term input compares to the total search volume. Another tool built around Google Search is Google Alerts, a content change detection tool that provides notiﬁcations automatically.

**Text analysis tools**

Text analysis tools are broad-based tools for natural language processing and text analysis. Examples of companies in the text analysis area include: OpenAmplify and Jodange whose tools automatically ﬁlter and aggregate thoughts, feelings and statements from traditional and social media. There are also a large number of freely available tools produced by academic groups and non-governmental organizations (NGO) for sourcing, searching and analyzing opinions. These include, but are not limited to,

* **Stanford NLP group tools and LingPipe** - a suite of Java libraries for the linguistic analysis of human language (Teuﬂ et al 2010). A variety of open-source text analytics tools are available, especially for sentiment analysis.
* **NLTK Natural Language Toolkit** - this tool includes open-source Python modules, linguistic data and documentation for text analytics. Another one is GATE (http:// gate.ac.uk/sentiment).
* **Lexalytics Sentiment Toolkit** – this tool performs automatic sentiment analysis on input documents. It is powerful when used on a large number of documents, but it does not perform data scraping.

**Data visualization tools**

The data visualization tools provide business intelligence (BI) capabilities and allow different types of users to gain insights from the ‘big’ data. The users can perform exploratory analysis through interactive user interfaces available on the majority of devices, with a recent focus on mobile devices). The data visualization tools help the users identify patterns, trends and relationships in the data which were previously latent. Fast ad hoc visualization on the data can reveal patterns and outliers, and it can be performed on large-scale data sets frameworks, such as Apache Hadoop or Amazon Kinesis. Two notable visualization tools are SAS Visual Analytics and Tableau.

**Existing Work**

In order to perform experiments, the scarcity of real data publically available and lack of properly researched methods and techniques publications are the two most often considered criticisms related to the research of suspicious detection based on data mining.

Table 2 below provides the briefs of various applications developed so far using the online or social media data.

**Table 2:** Various *applications developed till date using the online or social media data*

|  |  |
| --- | --- |
| **Author** | **Applications** |
| Ku, Liang et al., 2006  Schober, Pasek et al., 2016  Dodds, Harris et al., 2011  Smith 2002  Kempe, kleinberg et al.,2015 | Enable the tracking of online opinions in weblogs.  Highlights how participants view different elicitation techniques from surveys to social media data and the nature of the data  Measuring people’s happiness and mood.  Browsing document collections in order to detect events  Analyse the diffusion of ideas in social networks |

The past researches focused on mining facts but the recent research focuses on mining opinions. Detection of opinion from the web data deals with numerous things. (Liu, Hu et al., 2005, Mishne and Glance 2006). A recent paper on fraud detection (Phua, Lee et al., 2010), illustrates various categories of frauds identified by analyzing the web contents (listed in table ) for example frauds related to home insurance, crop insurance, automobile insurance and medical insurance.

In order to identify the general trends of these suspicious or fraudulent transactions and applications is mainly focused by these detection systems. Besides these domains, this research also identifies other probable areas where suspicious discussions are monitored using in text mining.

**Text mining**

Enormous amounts of messages get published each day on social media sites. For example, Twitter processes 230 million tweets (messages that are 140 characters long) a day (twitterstats). The explosion of textual messages can cause information overload. Our goal is to design systems that can analyze and summarize social media content. A basic process of text mining approach is shown in Figure below.

**Multiple sources, online, documents etc**

**Fig 2:** *Text mining process*

Recently the Facebook static messages are scanned to identify criminal’s behaviour. Also, in 2015, Siguenza-Guzman, et al. (2015) presents a literature review of data mining applications in academic libraries. In this they have identified various techniques to monitor special category of words required for a specific journal or library.

Diaz, et al. (2016) discusses online and social media data as an imperfect continuous panel survey. One more research article (Tayal, et al., 2015) identifies various data mining techniques for crime detection and criminal identification specifically in India.

**Study Gap**

In recent years, much research has focused on understanding the expression of individuals’ opinion online, and exploring its use as an alternative data collection modality for surveys and fundamental ways to gather the information and predict opinion, identify knowledge, support, and related tasks. However, finding accurate information in social big data is becoming a challenge for public and private research organization. Moreover, the evolution of internet has led to the growth of more innumerable cybercrimes.

Criminals use social networking websites, cell phones, messenger applications to send suspicious messages, thus making dynamically tracing their activities more difficult. Table below gives the particulars of work done by various authors during the span of time and the corresponding study gap.

Previously, the text mining techniques were based on brute force approaches, stemming algorithms, and keyword identification techniques. But due to the huge and continuous real time data, these techniques failed to achieve desired success rate. As a result, researchers shift their focus towards intelligent data and text mining approaches, which has the calibre to tell something about the world, outside the data collections themselves. Novel based investigation of the text-mining approaches should be based on the process of deliberately creating new knowledge that was not existed before and cannot be simply discovered or retrieved. The process is manual in nature i.e. mostly done by humans only and cannot be automated easily.

Artificial Intelligence can be used to facilitate this novel investigation as it simulates human behaviour and intelligence. Such an automatic system could be called either as intelligent test mining (for considering the structured textual data) or intelligent data mining (for considering structured data in the form alphanumeric and numeric fields.

Keeping this in mind, the present research work will consider the swarm intelligence mechanisms (subcategory of soft computing and artificial intelligence domain) to design a data mining technique which will efficiently and effectively monitor the suspicious discussions on the social media websites’ data. This will surely provide an upper hand to the law enforcement agencies throughout the world, which are in ultimate need of these kinds of systems to have prior knowledge about the crimes.

**Table 3:** *Study gap and work done by various authors*

|  |  |  |  |
| --- | --- | --- | --- |
| **Author & Year** | **Research** | **Worked On** | **Gap** |
| **Alami and Elbequali, 2015**  **Murugesun et al, 2016**  **Alami and Elbequali, 2015** | Cybercrime Profiling: Text mining technique to predict and detect criminal activities  Automated Monitoring of Suspicious Discussions on Online Forums using Corpus based approach  Detecting Suspicious profiles using text analysis within social media | hashtags on Twitter (e.g., #arabspring, #BostonAttack) were used especially to target and detect suspicious topics and eventual illegal events. Similarity approach is used in text analysis to detect suspicious posts  Monitor discussion for illegal activities and download postings that are in texts format as evidence for investigation.  Text analysis posted generally by social media users in order to discover the suspicious published content with the deduction of suspicious behaviour users on social media | Based on predefined suspicious words. Resulting in limited text corpus database\  Based on the traditional approach of similarity matching between the words  Using similarity difference leading to high execution time and lower precision |

**Chapter 3: Methodology**

**Introduction**

This chapter addresses the methods, procedures and instruments to be used by the researcher to gather data and analyze them. The chapter covers the development methodologies that were used in the study.

**Social media methodology**

The two major impediments to using social media for academic research are ﬁrstly access to comprehensive data sets and secondly tools that allow ‘deep’ data analysis without the need to be able to program in a language such as Java. The majority of social media resources are commercial and companies are naturally trying to monetize their data. As discussed, it is important that researchers have access to open-source ‘big’ (social media) data sets and facilities for experimentation. Otherwise, social media research could become the exclusive domain of major companies, government agencies and a privileged set of academic researchers presiding over private data from which they produce papers that cannot be critiqued or replicated. Recently, there has been a modest response, as Twitter and Gnip are piloting a new program for data access, starting with 5 all-access data grants to select applicants.

**Software Development Life Cycle Methodology**

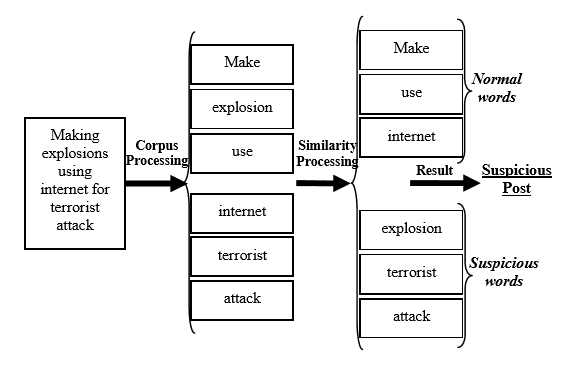
“Agile software development describes a set of principles for software development under which requirements and solutions evolve through the collaborative effort of self-organizing crossfunctional teams” (Collier, 2011). Larman (2004) demonstrated agile software development as the promotion of adaptive planning, development of evolutionary, early delivery, and continual improvement of software while encouraging rapid and flexible response to change.

Agile software development defines software development as a group of methodologies that promote iterations, open collaborating and adaptability of processes throughout the life-cycle of the development of the project. It allows the developer to perform small increments with minimal planning on a, develop as you go basis, rather than planning the whole development at length. This minimizes the overall risk and encourages the project in adapting to changes more frequently. It places an emphasis on stakeholder and code developer involvement translating to the stakeholder being consulted about the product iterations and comments noted as software development proceeds.

This methodology was employed for the development of the web interface, database, search query components and sentiment analysis tools for illegal behaviour ranking. Documentation, requirements documents and design documents, is a key part of the methodology as well as the source code. This encourages that the software development continues even if one or more members of the development team leaves the whole project does not collapse. The complete working design document is important as new team members or a completely different new team can be easily brought to speed on the progress of the design and easily familiarize themselves when undertaking the software development (Hughey, 2009).

**Design Approach**

For sake of research and experimental purposes we collect data on four different categories namely; Political, Sexual, Radical, Violence and Criminal. All the data about each category are collected from Twitter and Facebook.

**

**Fig:** *Detection of suspicious posts using a similarity approach*

**Data Sources**

All the data received on the Internet in Zimbabwe would have been the appropriate sample specimen. However, because of the numerous daily Internet data, forums, websites and discussion chat. The researcher was to target the popular website and web forums. Twitter and Facebook were to be considered to participate in the dark web collection of suspicious posts and data. A set of keywords such as ‘Kill’, ‘Burn’, ‘Uraya’,’Rovai’ will be used as search tools to query the database of the collected social media network.

**Data Collection**

With the data sources identified a means for data collection, harvesting, cleaning and verification was implemented. The study mainly focuses on social media data from websites, Facebook, Twitter and blogs. Harvesting data from social media is provided for by open source and several APIs which are specific to the social media site. The researcher used APIs as they are specific and proprietary. Python was the progamming language that was used to perform web crawling and scraping. The use of Python was chosen due to its versatility, agility and previous studies have shown it to be a viable solution for web crawling, spidering, indexing and scraping. Labels of positive, negative and neutral are appended to the training set of social media data. The collected data includes text, emoticons and common acronyms.

An application designed in any dedicated web-programming language like Java, PHP, Python could be developed to access the APIs of these networks. Characteristics of data extraction methodologies for some famous networking platforms are describes below.

* **Facebook -** Facebook data are accessible through different methods that can be summed up into two main branches Graph API and FQL (Facebook Query Language).

*API description -* The Graph API is a way that tries to simplify the vision of each object contained in Facebook. The idea is that a user, a picture, a comment, a page are regarded as single elements in the platform, connected by some particular relations (e.g. the list of photo elements in an album of a user). The Graph API is very useful for gathering individual data of an object. FQL, based on the Graph API, allows to make queries on the platform for gathering bunches of data and turns out to be more readable in some ways when it comes to obtain the data you want.

*Restriction of Data Request -* Data access need permission to fetch non public data: nowadays Facebook has about 60 different permission levels. Facebook has a rate limit of 600 calls/600 minutes.

* **Twitter** - Twitter data are accessible through Tweepy

*API description -* The REST API enables developers to access some of the core primitives of Twitter including timelines, status updates, and user information. In addition to offering programmatic access to the timeline, status, and user objects, this API also enables developers a multitude of integration opportunities to interact with Twitter. The Streaming API is the real-time sample of the Twitter Firehose. This API is for those developers with data intensive needs. The Streaming API is most suited for data mining and analytics research. Streaming API allows for large quantities of key

*Restriction of Data Request -* Authentication is required to access the data by API methods. The data retrieval from twitter is restricted by rate limiting: data calls are permitted 150 requests per hour. Words to be specified and tracked, retrieving geotagged tweets from a certain region, or have the public statuses of a user set returned.

**Social networking media**

As with Wikipedia, popular social networks, such as Facebook, Twitter and Foursquare, make a proportion of their data accessible via APIs. Although many social networking media sites provide APIs, not all sites (e.g., Bing, LinkedIn and Skype) provide API access for scraping data. While more and more social networks are shifting to publicly available content, many leading networks are restricting free access, even to academics.

**Twitter Data Collection**

Twitter is a popular social media place to look for information about literally any product. Also the nature of tweets in twitter that its length cannot be greater than 140 characters interests us because we deal with sentence level sentiment detection in the following sections. Twitter provides an open source library to access the tweets programmatically, there are many wrappers developed and in our case we use the twitter4j library. The API (application program interface) allows inputting a search query along with various other preferences like;

* Location of the tweet,
* Language of the tweet, (the researcher interested in only English tweets since we do not focus on detecting sentiments of other languages),
* Username of the person who tweeted, but we do not take in this information.
* Tweets between certain period, or tweets since a date or tweets until a date. Twitter does not provide any data older than a week hence it is necessary to run the data retrieval program continuously in order to obtain updated tweets.
* Re-tweets, we can fetch the replies to a tweet also.
* Attitudes - Twitter has inbuilt program that classifies a tweet into positive, negative and neutral attitudes, pretty much the same the researcher is attempting to do.

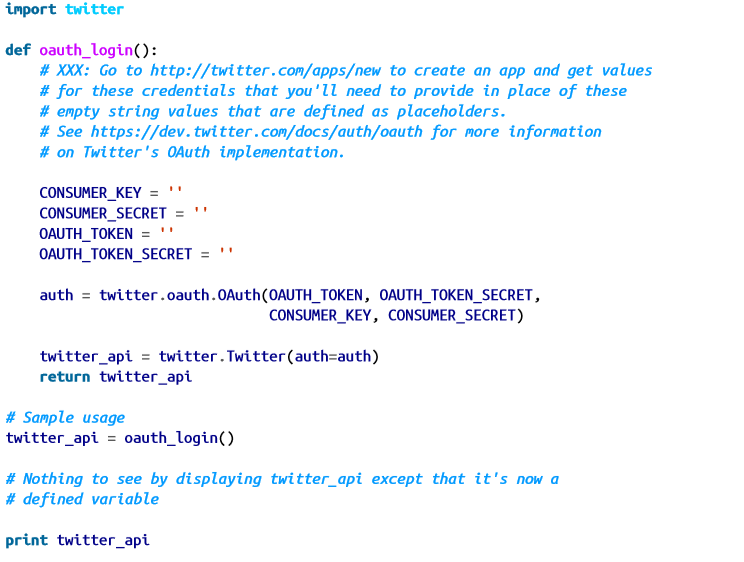
With all the above features we can collect tweets of our choice. It was possible to collect the comments with all the features except attitude using our program. It was possible to collect 100 comments per run and managed to collect more comments in a real time basis. The researcher could both search Twitter and listen to Twitter streams (that is print out Tweets in real-time that match a particular query).

The default account setting keeps users’ Tweets public, although users can protect their Tweets and make them visible only to their approved Twitter followers. However, less than 10 % of all the Twitter accounts are private. Tweets from public accounts (including replies and mentions) are available in JSON format through Twitter’s Search API for batch requests of past data and Streaming API for near real-time data.

* **Search API** - Query Twitter for recent Tweets containing speciﬁc keywords. It is part of the Twitter REST API v1.1 (it attempts to comply with the design principles of the REST architectural style, which stands for Representational State Transfer) and requires an authorized application (using oAuth, the open standard for authorization) before retrieving any results from the API.
* **Streaming API** - A real-time stream of Tweets, ﬁltered by user ID, keyword, geographic location or random sampling.

The researcher retrieved recent Tweets containing particular keywords through Twitter’s Search API Twitter’s Streaming API allows data to be accessed via ﬁltering (by keywords, user IDs or location) or by sampling of all updates from a select amount of users. Default access level ‘Spritzer’ allows sampling of roughly 1 % of all public statuses, with the option to retrieve 10 % of all statuses via the ‘Gardenhose’ access level (more suitable for data mining and research applications).

In social media, streaming APIs are often called Firehose, a syndication feed that publishes all public activities as they happen in one big stream. Twitter results are stored in a JSON array of objects containing the ﬁelds. The JSON array consists of a list of objects matching the supplied ﬁlters and the search string, where each object is a Tweet and its structure is clearly speciﬁed by the object’s ﬁelds, e.g., ‘created\_at’ and ‘from user’.



**Fig:** *Authorizing an application to access Twitter account data*

**Facebook Data Collection**

Facebook’s privacy issues are more complex than Twitter’s, meaning that a lot of status messages are harder to obtain than Tweets, requiring ‘open authorization’ status from users. Facebook currently stores all data as objects1 and has a series of APIs, ranging from the Graph and Public Feed APIs to Keyword Insight API. In order to access the properties of an object, its unique ID must be known to make the API call.

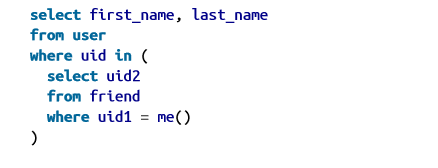
**Fig:** *Facebook applications must explicitly request authorization to access a user’s account data. Top: The Graph API Explorer permissions panel.*

* **Facebook’s Search API** (part of Facebook’s Graph API) can be accessed by calling https://graph.facebook.com/search?q= QUERY&type=page. ‘QUERY’ can be replaced by any search term, and ‘page’ can be replaced with ‘post,’ ‘user,’ ‘page,’ ‘event,’ ‘group,’ ‘place,’ ‘checkin,’ ‘location’ or ‘placetopic.’ The results of this search will contain the unique ID for each object. When returning the individual ID for a particular search result, one can use https://graph. facebook.com/ID to obtain further page details such as number of ‘likes.’
* **Facebook Graph API** search queries require an access token included in the request. Searching for pages and places requires an ‘app access token’, whereas searching for other types requires a user access token. Replacing ‘page’ with ‘post’ in the aforementioned search URL will return all public statuses containing this search term.

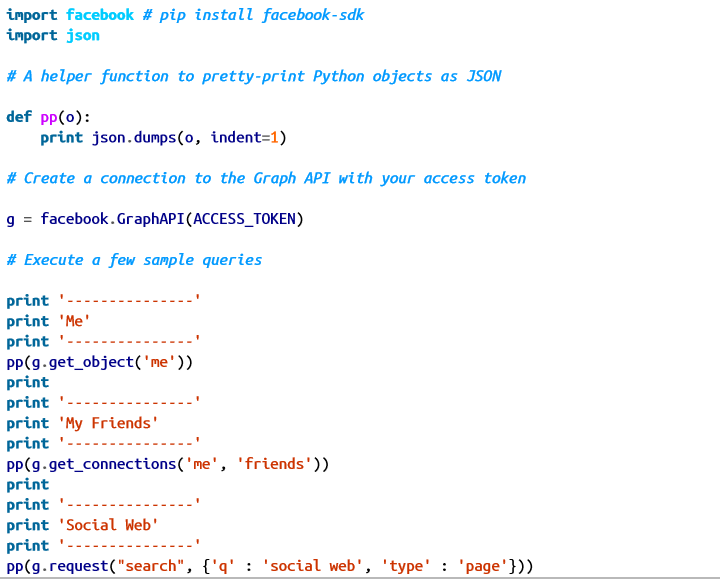
**Fig:** *A Facebook dialog requesting authorization for the Graph API Explorer application to access pages’ posts data*

**Facebook Query Language (FQL)**

In addition to the Graph API, FQL provides a fine alternative fpr querying Facebook’s Social Graph and has a SQL-inspired syntax that most developers find intuitive. It seems to be the case that any data you could query with the Graph API, you could also query via FQL, and although it may be true that some advanced queries that are possible with FQL may not be possible with the Graph API, it appears that Facebook’s longer-term plan is to ensure that the Graph API is at full parity with FQL. For example, you could query the first and last names of your friends in the FQL Query tab of the Graph API Explorer with the following FQL query:



**Querying the Graph API with Python**

****

Facebook also returns data in JSON format and so can be retrieved and stored using the same methods as used with data from Twitter, although the ﬁelds are different depending on the search type. The interaction between Pattern and Facebook is a bit more complex since you actually have to use an access key to allow Pattern to access your Facebook data through Python.

**Data Preparation and Transformation**

The opinion mined data is usually unstructured and contains irrelevant and non-textual characters and thus requires to be prepared, processed and transformed for data evaluation and validation. Text preparation involves cleaning before analysis is performed (Rambocas et al, 2013).

**Natural Language Processing on Social Media**

NLP is a long established sub-field of artificial intelligence research. It combines approaches developed in the fields of computer science, applied mathematics and linguistics. It ‘teaches’ algorithms to automatically detect the meaning of ‘natural’ language, such as that found on social media. These algorithmic models look for statistical correlations between the language used and the meaning expressed on the basis of previous examples provided by human analysts, and then, building on this, automatically (and therefore at great speed) make decisions about the meaning of additional, unseen messages. NLP is increasingly and necessarily used as an analytical ‘window’ into datasets of social media communication that are too large to be manually analysed.

This training of NLP algorithms, a technique called Machine Learning, is conducted through a process called ‘mark up’. Messages are presented to the analyst via an interface. The analyst reads each message, and decides which of a number of pre-assigned categories of meaning it best fits. After the analyst has made a decision, they click on the most relevant tweet and it is ‘annotated’, becoming associated with that category. The NLP algorithm then appraises the linguistic attributes, that, depending on the specific algorithm, often includes words (or unigrams), collection of words (such as bigrams and trigrams), grammar, word order or emoticons – that correlate strongly with each category. These measured correlations provide the criteria for which the algorithm then proceeds to make additional automatic judgments about which category additional (and un-annotated) pieces of social media data best fit into.

The statistical nature of this approach renders it notionally applicable to any language where there is a statistical correlation between language use and meaning. NLP programmes vary in the way they make their decisions: some place more weight on specific words, others on structural or grammatical features.

The operational opportunity of NLP for countering terrorism is to use these algorithmic models as ‘classifiers’. Classifiers are applied NLP algorithms that are trained to categorise each piece of social media data (each post or tweet) into one of a small number of pre-defined categories. The earliest and most widely applied example of this technology is ‘sentiment analysis’, wherein classifiers make decisions on whether a piece of social media data is broadly positive or negative in tone. However, the kinds of distinctions that a classifier can make are arbitrary, and can be determined by the analyst and the context.

The performance of NLP classifiers is often quantified by comparing a body of automatically classified data against a ‘gold standard’ set of human classifications. On this measure, their accuracy, the ability of the NLP algorithm to classify any given message the same way a human would, varies considerably. There are many scenarios where 90 per cent accuracy would be expected. However, an accuracy of around 70-80 per cent in a three-way classification task would often be considered excellent.

Classifiers are sensitive to the specific vocabulary seen in the data used to train them. The best classifiers are therefore also highly bespoke and trained on a specific conversation at a specific time to understand context-specific significance and meaning. As language use and meaning constantly change, the classifier must be retrained to maintain these levels of accuracy. The more generic and expansive the use of any NLP classifier, the more likely that it will misunderstand language use, misclassify text and return inaccurate results.

In many situations, the performance of these classifiers is sufficient to produce robust aggregate findings, even when the accuracy of any given singular classification is quite low. This arises because the data sets are sufficiently large that even relatively inaccurate individual estimates lead to an accurate assessment of the overall trend. Assessing when this technology tends to work well and when it does not is an area of active research.

A key area of active research is in the reduction of the time, effort and cost required to train and maintain an NLP classifier. It is typically very expensive to produce the labeled training data that these supervised machine learning algorithms require. In complex tasks, it would not be unusual for this to take multiple person months of effort. The novel introduction of an information theoretic technique called ‘active learning’ is beginning to allow classifiers to be built much more rapidly and cheaply often in a matter of hours, and sufficiently quickly to meet changing operational requirements prompted by rapidly shifting situations and contexts.

*There are three emerging uses of NLP that the researcher considered particularly relevant.*

1. The **first** is to classify tweets into categories other than positive, negative and neutral: such as urgent, calm, violent or pacific.
2. The **second** is to use NLP to dramatically reduce the amount of data that an analyst must sift through in order to find messages of relevance or interest. In this respect, classifiers can also be 'tuned' to perform at high precision (only highlighting messages very likely to be of interest) or high recall (highlighting all messages conceivably of interest). This form of relevancy filtering is sometimes known as ‘disambiguation’.
3. The **third** is to create layers of multiple NLP classifiers to make architectures capable of making more sophisticated decisions.

**Natural Language Toolkit (NTLK)**

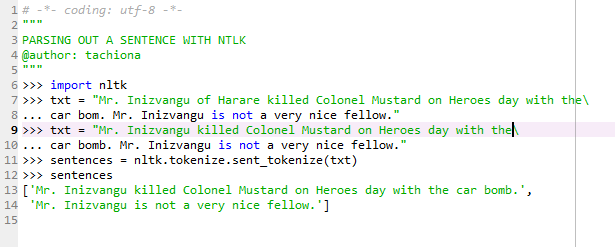
NLTK is written such that you can explore data easily and begin to form some impressions without a lot of upfront investment.

**Natural Language Processing Illustrated Step-by-Step**

The researcher prepared to step through a series of examples that illustrate NLP with NLTK. The NLP pipeline we’ll examine involves the following steps:

**Step One: EOS detection**

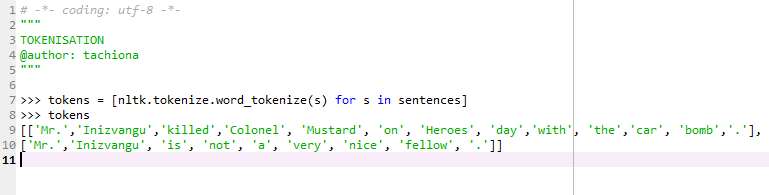
This step breaks a text into a collection of meaningful sentences. Since sentences generally represent logical units of thought, they tend to have a predictable syntax that lends itself well to further analysis.

****

**Fig:** *Parsing out a sentence with NTLK*

**Step Two: Tokenization**

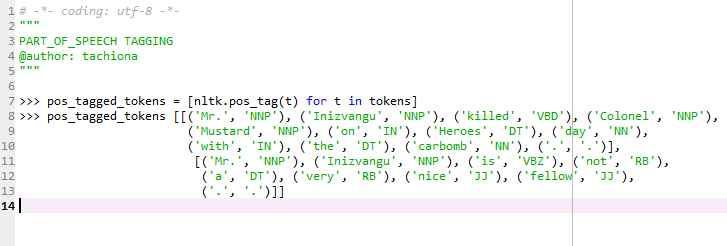
This step operates on individual sentences, splitting them into tokens. The researcher notes that distinguishing between whether a period is an EOS marker or part of an abbreviation isn’t always trivial. As an anecdotal note, some written languages, such as ones that use pictograms as opposed to letters, don’t necessarily even require whitespace to separate the tokens in sentences and require the reader (or machine) to distinguish the boundaries.



**Fig:** *Tokenization process*

**Step Three: POS tagging**

This step assigns part-of-speech (POS) information to each token.



**Fig:** *Part-Of-Speech tagging*

**Step Four: Chunking**

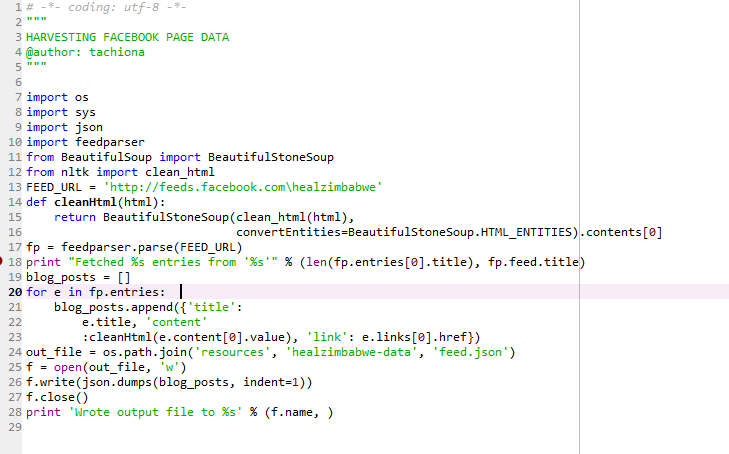
This step involves analyzing each tagged token within a sentence and assembling compound tokens that express logical concepts quite a different approach than statistically analyzing collocations. It is possible to define a custom grammar through NLTK’s chunk.

**Step Four: Extraction**

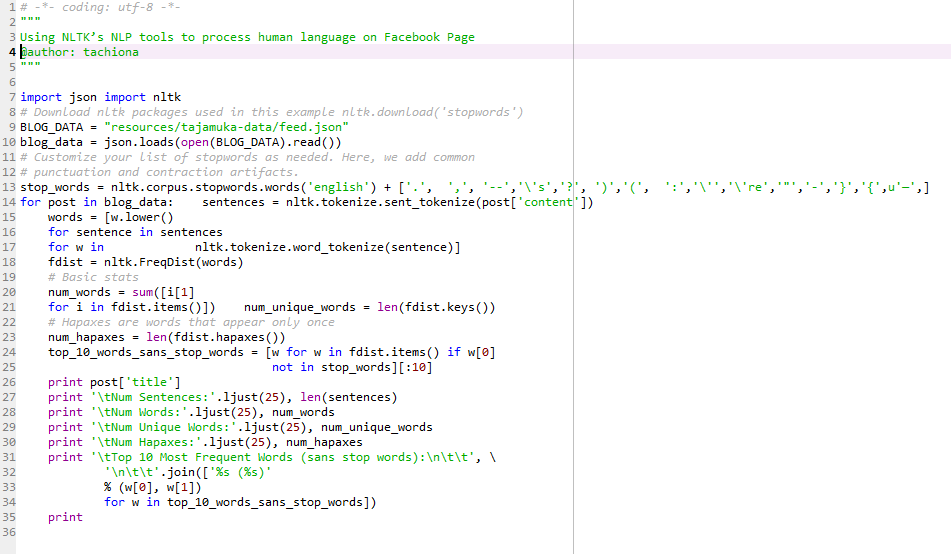
This step involves analyzing each chunk and further tagging the chunks as named entities, such as people, organizations, locations, etc.

**Sentence Detection in Human Language Data**

The researcher found out that EOS detection alone yields some powerful possibilities, such as document summarization. But first, the researcher needed to fetch some clean human language data. The researcher used the tried-and-true feed parser package, along with utilities that are based on nltk and BeautifulSoup, to clean up HTML formatting that may appear in the content to fetch some posts from Facebook and Twitter.

****

**Fig:** *Harvesting social media data by parsing feeds*

****

**Fig:** *Using NLTK’s NLP tools to process human language in social media data*

**Social Media Data Analysis**

In the research great lengths have been made to assess the data to the best it can be done. Automatic filtering and analysis has been implemented so as to reduce the data set and making continual assessment and monitoring of suspicious activities. The manual data filtering and analysis will seek to seek to provide a contextual lookup of the data that has been flagged as suspicious

As we are aware that with words and information posted in social media, there is a need to understand the context in which the word was used. For example, bomb is a potential word for flagging down as suspicious but if used in a different context it can be an affirmation. Such as in the case a friend tells a friend you are the bomb, meaning that the other friend is a cool and nice friend.

Hence, the need for a semi-automated tool is critical as it provides the constant automatic monitoring and ease of work. But also the human knowledge element is important to remove the clutter and jargon flagged as suspicious by the system.

The data collected from the Facebook and Twitter accounts through the APIs will form the database. The raw data will undergo filtering by using the keywords as search tools which will be run as queries on the database. Once the filtered data is collected it will be mapped against suspicious user together with the total counts of suspicious words and exact phrase of suspicious post.

The qualitative analysis will be undertaken by the Cyber Forensic expert to check if the conversation was taken out of context. Once the quantitative and qualitative analysis is complete the data is confirmed as the true reflection of the data.

**Summary**

This research thesis target was to investigate the social media web forums for suspicious behaviour and report and blacklist them. Additionally, it was to provide the forensic digital evidence. The research was intended to collect dark Facebook and Twitter posts and from it mine the opinions that are suspicious and classify them.

The results from the inputs would pave the way for the development of the system and framework for collecting suspicious posts, mining suspicious opinions from the data collected and to provide a systematic approach of the digital evidence collected.

**Chapter 4: Results and Discussion**

**Introduction**

This chapter presents and discusses the results obtained from study.

**Presentation of Results**

The Social Media Forensics was designed and developed to accept the URL and perform the digital forensics of copying the data and placing it in a database.

**Twitter Results and Discussion**

**Scenario One:** *Discovering the Trending Topics on Twitter*

**Problem:** The researcher wanted to know what is trending on Twitter for a particular geographic area such as the Zimbabwe, another country or group of countries, or possibly even the entire world.

**Solution:** Twitter’s Trends API enabled the researcher to get the trending topics for geographic areas that are designated by a Where On Earth (WOE) ID, as defined and maintained by Yahoo!

**Discussion:** A place is an essential concept in Twitter’s development platform, and trending topics are accordingly constrained by geography to provide the best API possible for querying for trending topics. Like all other APIs, it returns the trending topics as JSON data, which can be converted to standard Python objects and then manipulated with list comprehensions or similar techniques. This means it’s fairly easy to explore the API responses.

**Sample results:**

**Scenario Two:** *Searching for Tweets*

**Problem:** The researcher needed to search Twitter for tweets using specific keywords and query constraints.

**Solution:** The researcher used the Search API to perform a custom query.

**Discussion:** Figure below illustrates how to use the Search API to perform a custom query against the entire Twitterverse. Similar to the way that search engines work, Twitter’s Search API returns results in batches, and you can configure the number of results per batch to a maximum value of 200 by using the count keyword parameter. It is possible that more than 200 results (or the maximum value that you specify for count) may be available for any given query, and in the parlance of Twitter’s API, we’ll need to use a cursor to navigate to the next batch of results. Cursors are a new enhancement to Twitter’s v1.1 API and provide a more robust scheme than the pagination paradigm offered by the v1.0 API, which involved specifying a page number and a results per page constraint. The essence of the cursor paradigm is that it is able to better accommodate the dynamic and real-time nature of the Twitter platform.

**Sample results:**

**Scenario Twelve:** *Finding Tweets Using a Keyword*

**Problem:** Getting the most recent tweets that contain a keyword. This can be extremely useful if you want to monitor specifically mentioned topics in the Twitter world. The researcher wanted to see how Twitter’s been mentioning Coup in Zimbabwe.

**Solution:** The researcher employed the search() function as it was the best tool to accomplish the goal. The most important parameter here is q, the query parameter, which is the keyword we’re searching for. The researcher also set the language parameter so we don’t get any tweets from an unwanted language and only returned English (“en”) tweets.

**Sample results:**

**Scenario Three:** *Analyzing Tweet Content*

**Problem:** Given a collection of tweets, the researcher wanted to do some cursory analysis of the 140 characters of content in each to get a better idea of the nature of discussion and ideas being conveyed in the tweets themselves.

**Solution:** The researcher used simple statistics, such as lexical diversity and average number of words per tweet, to gain elementary insight into what is being talked about as a first step in sizing up the nature of the language being used.

**Discussion:** In addition to analyzing the content for tweet entities and conducting simple frequency analysis of commonly occurring words, you can also examine the lexical diversity of the tweets and calculate other simple statistics, such as the average number of words per tweet, to better size up the data. Lexical diversity is a simple statistic that is defined as the number of unique words divided by the number of total words in a corpus; by definition, a lexical diversity of 1.0 would mean that all words in a corpus were unique, while a lexical diversity that approaches 0.0 implies more duplicate words. Depending on the context, lexical diversity can be interpreted slightly differently.

In the Twittersphere, lexical diversity might be interpreted in a similar fashion if comparing two Twitter users, but it might also suggest a lot about the relative diversity of overall content being discussed, as might be the case with someone who talks only about technology versus someone who talks about a much wider range of topics. In a context such as a collection of tweets by multiple authors about the same topic (as would be the case in examining a collection of tweets returned by the Search API or the Streaming API), a much lower than expected lexical diversity might also imply that there is a lot of “group think” going on. Another possibility is a lot of retweeting, in which the same information is more or less being regurgitated. As with any other analysis, no statistic should be interpreted devoid of supporting context.

**Sample Results:**

**Scenario Four:** *Analyzing a User’s Favorite Tweets*

**Problem:** The researcher needed to learn more about what a person cares about by examining the tweets that a person has marked as favorites.

**Solution:** The researcher used the GET favorites/list API endpoint to fetch a user’s favorite tweets and then apply techniques to detect, extract, and count tweet entities to characterize the content.

**Discussion:** Not all Twitter users take advantage of the bookmarking feature to identify favorites, so you can’t consider it a completely dependable technique for zeroing in on content and topics of interest; however, if you are fortunate enough to encounter a Twitter user who tends to bookmark favorites as a habit, you’ll often find a treasure trove of curated content. In addition to favorites, any tweets that a user has retweeted are also promising candidates for analysis, and even analyzing patterns of behavior such as whether or not a user tends to retweet (and how often), bookmark (and how often), or both is an enlightening survey in its own right.

**Sample Results:**

**Scenario Five:** *Extracting Tweet Entities*

**Problem:** The researcher wanted to extract entities such as @username mentions, #hashtags, and URLs from tweets for analysis.

**Solution:** Extract the tweet entities from the entities field of tweets.

**Discussion:** Twitter’s API now provides tweet entities as a standard field for most of its API responses, where applicable. The entities field, illustrated in figure below, includes user mentions, hashtags, references to URLs, media objects (such as images and videos), and financial symbols such as stock tickers. At the current time, not all fields may apply for all situations. For example, the media field will appear and be populated in a tweet only if a user embeds the media using a Twitter client that specifically uses a particular API for embedding the content; simply copying/pasting a link to a YouTube video won’t necessarily populate this field.

**Sample results:**

**Scenario Six:** *Finding Users Who Have Retweeted a Status*

**Problem:** The researcher sought to discover all of the users who have ever retweeted a particular status.

**Solution:** The researcher used the GET retweeters/ids API endpoint to determine which users have retweeted the status.

**Discussion:** Although the GET retweeters/ids API returns the IDs of any users who have retweeted a status, there are a couple of subtle caveats that you should know about. In particular, keep in mind that this API reports only users who have retweeted by using Twitter’s native retweet API, as opposed to users who have copy/pasted a tweet and prepended it with “RT,” appended attribution with “(via @exampleUser),” or used another common convention. Most Twitter applications (including the twitter.com user interface) use the native retweet API, but some users may still elect to share a status by “working around” the native API for the purposes of attaching additional commentary to a tweet or inserting themselves into a conversation that they’d otherwise be broadcasting only as an intermediary.

**Sample results:**

**Scenario Seven:** *Resolving User Profile Information*

**Problem:** The researcher to look up profile information for one or more user IDs or screen names.

**Solution:** The researcher used the GET users/lookup API to exchange as many as 100 IDs or usernames at a time for complete user profiles.

**Discussion:** Many APIs, such as GET friends/ids and GET followers/ids, return opaque ID values that need to be resolved to usernames or other profile information for meaningful analysis. Twitter provides a GET users/lookup API that can be used to resolve as many as 100 IDs or usernames at a time, and a simple pattern can be employed to iterate over larger batches. Although it adds a little bit of complexity to the logic, a single function can be constructed that accepts keyword parameters for your choice of either usernames or IDs that are resolved to user profiles.

**Sample results:**

**Scenario Eight:** *Getting All Friends or Followers for a User*

**Problem:** The researcher sought to harvest all of the friends or followers for a (potentially very popular) Twitter user.

**Solution:** The researcher used the make\_twitter\_request function to simplify the process of harvesting IDs by accounting for situations in which the number of followers may exceed what can be fetched within the prescribed rate limits.

**Discussion:** The GET followers/ids and GET friends/ids provide an API that can be navigated to retrieve all of the follower and friend IDs for a particular user, but the logic involved in retrieving all of the IDs can be nontrivial since each API request returns at most 5,000 IDs at a time. Although most users won’t have anywhere near 5,000 friends or followers, some celebrity users, who are often interesting to analyze, will have hundreds of thousands or even millions of followers. Harvesting all of these IDs can be challenging because of the need to walk the cursor for each batch of results and also account for possible HTTP errors along the way. Fortunately, it’s not too difficult to adapt make\_twitter\_request and previously introduced logic for walking the cursor of results to systematically fetch all of these ids.

**Sample results:**

**Scenario Nine:** *Analyzing a User’s Friends and Followers*

**Problem:** The researcher sought to conduct a basic analysis that compares a user’s friends and followers.

**Solution:** The researcher used the setwise operations such as intersection and difference to analyze the user’s friends and followers.

**Discussion:** After harvesting all of a user’s friends and followers, you can conduct some primitive analyses using only the ID values themselves with the help of setwise operations such as intersection and difference.

**Sample results:**

**Scenario Ten:** *Harvesting a User’s Tweets*

**Problem:** The researcher wanted to harvest all of a user’s most recent tweets for analysis.

**Solution:** The author used the GET statuses/user\_timeline API endpoint to retrieve as many as 3,200 of the most recent tweets from a user, preferably with the added help of a robust API wrapper such as make\_twitter\_request since this series of requests may exceed rate limits or encounter HTTP errors along the way.

**Discussion:** Timelines are a fundamental concept in the Twitter developer ecosystem, and Twitter provides a convenient API endpoint for the purpose of harvesting tweets by user through the concept of a “user timeline.” Harvesting a user’s tweets, is a meaningful starting point for analysis since a tweet is the most fundamental primitive in the ecosystem. A large collection of tweets by a particular user provides an incredible amount of insight into what the person talks (and thus cares) about. With an archive of several hundred tweets for a particular user, you can conduct dozens of experiments, often with little additional API access. Storing the tweets in a particular collection of a document-oriented database such as MongoDB is a natural way to store and access the data during experimentation. For longer-term Twitter users, performing a time series analysis of how interests or sentiments have changed over time might be a worthwhile exercise.

**Sample results:**

**Scenario Eleven:** *Crawling a Friendship Graph*

**Problem:** The author sought to harvest the IDs of a user’s followers, followers of those followers, followers of followers of those followers, and so on, as part of a network analysis essentially crawling a friendship graph of the “following” relationships on Twitter.

**Solution:** The author used a breadth-first search to systematically harvest friendship information that can rather easily be interpreted as a graph for network analysis.

**Discussion:** A breadth-first search is a common technique for exploring a graph and is one of the standard ways that you would start at a point and build up multiple layers of context defined by relationships. Given a starting point and a depth, a breadth-first traversal systematically explores the space such that it is guaranteed to eventually return all nodes in the graph up to the said depth, and the search explores the space such that each depth completes before the next depth is begun.

**Sample results:**

**Facebook Monitoring Results and Discussions**

Facebook is arguably the heart of the social web and is somewhat of an all-in-one wonder, given that more than half of its 1 billion users1 are active each day updating statuses, posting photos, exchanging messages, chatting in real time, checking in to physical locales, playing games, shopping, and just about anything else you can imagine. From a social web mining standpoint, the wealth of data that Facebook stores about individuals, groups, and products is quite exciting, because Facebook’s clean API presents incredible opportunities to synthesize it into information (the world’s most precious commodity), and glean valuable insights. On the other hand, this great power commands great responsibility, and Facebook has instrumented the most sophisticated set of online privacy controls that the world has ever seen in order to help protect its users from exploit. It’s worth noting that although Facebook is self-proclaimed as a social graph, it’s been steadily transforming into a valuable interest graph as well, because it maintains relationships between people and the things that they’re interested in through its Facebook pages and “Likes” feature.

**Understanding the Social Graph API**

As its name implies, Facebook’s Social Graph is a massive graph data structure representing social interactions and consisting of nodes and connections between the nodes. The Graph API provides the primary means of interacting with the Social Graph, and the best way to get acquainted with the Graph API is to spend a few minutes tinkering around with the Graph API Explorer. It is important to note that the Graph API Explorer is not a particularly special tool of any kind. Aside from being able to prepopulate and debug your access token, it is an ordinary Facebook app that uses the same developer APIs that any other developer application would use. The results of a Graph API query are returned in a convenient JSON format that can be easily manipulated and processed.

**Fig:** *Using the Graph API Explorer application to progressively build up a query for users’ interests and posts: a query for a node in the Social Graph*

**Fig:** *Using the Graph API Explorer application to progressively build up a query for users’ interests: a query for a node, connections to friends, followers and comments and likes for those friends*

**Analyzing Social Graph Connections**

The researcher used an official Python SDK for the Graph API is a community fork of that repository previously maintained by Facebook and can be installed per the standard protocol with pip via pip install facebook-sdk. This package contains a few useful convenience methods that allow you to interact with Facebook in a number of ways, including the ability to make FQL queries and post statuses or photos. The methods are:

1. get\_object(self, id, \*\*args)

*Example usage:* get\_object("me", metadata=1)

1. get\_objects(self, id, \*\*args)

*Example usage:* get\_objects(["me", "some\_other\_id"], metadata=1)

1. get\_connections(self, id, connection\_name, \*\*args)

*Example usage:* get\_connections("me", "friends", metadata=1) request(self, path,

1. args=None, post\_args=None)

*Example usage:* request("search", {"q" : "social web", "type" : "page"})

**Fig:** *Querying the Graph API with Python*

**Analyzing Facebook Pages**

Although Facebook started out as more of a pure social networking site without a Social Graph or a good way for businesses and other entities to have a presence, it quickly adapted to take advantage of the market needs. Fast-forward a few years, and now businesses, clubs, books, and many other kinds of nonperson entities have Facebook pages with a fan base.

Facebook pages are a powerful tool for businesses to engage their customers, and Facebook has gone to some lengths to provide tools that allow Facebook page administrators to understand their fans with a small toolbox that is appropriately called “Insights.” If you’re already a Facebook user, the chances are pretty good that you’ve already liked one or more Facebook pages that represent something that you approve of or think is interesting, and in this regard, Facebook pages significantly broaden the possibilities for the Social Graph as a platform. The explicit accommodation of nonperson user entities through Facebook pages, the Like button, and the Social Graph fabric collectively provide a powerful arsenal for an interest graph platform, which carries with it a profundity of possibilities.

**Examining Friendships and Connections**

The researcher made use of knowledge of the Graph API to examine the friendships from user’s network

* Are there any suspicious topics or special interests that are especially pronounced within the network?
* Does the suspicious user’s social network contain many mutual friendships or even larger cliques?
* How well connected are the people in discussing a suspicious topic?
* Are any of the friends particularly outspoken or passionate about suspicious posts?

The following illustrations show analyzing posts, comments and likes as well as analyzing and visualizing mutual friendships.

**Fig:** *Querying for all users’ posts, comments and likes*

**Fig:** *Calculating the most popular posts and topics of a user*

**Fig:** *Calculating the most popular categories for posts of a user*

**Fig:** *Finding common posts between an suspicious user and its friendships.*

**Fig:** *A graph of suspicious users’ mutual friendships within a Facebook social network*

**Summary**

**Chapter 5: Conclusions and Recommendations**

**Achievements**

The objectives of the study were discussed and the below were the procedures used to achieve them:

* To identify and analyze techniques used in hate speech monitoring and select the best suitable technique for creating a customized hate speech application. To achieve this objective, different techniques for hate speech monitoring were analyzed which included hate speech using machine language, natural processing algorithms and classifiers
* To demonstrate and test the application while providing analysis on the hate speech websites being investigated. A set of keywords that were used to search the stored crawled websites formed the basis for analysis. The frequency of the potential hate speech keywords was used to plot and quantify the sentiment of the web forum. This was represented statistically as shown on the table below where the application sentiment analysis was evaluated and cross-validation performed on the premise of testing the precision, recall, F1 score and support with the average accuracy score of the sentiment analysis tool as 75.89%.

**Evaluation**

As discussed, the easy availability of APIs provided by Twitter and Facebook has led to an ‘explosion’ of data services and software tools for scraping and sentiment analysis, and social media analytics platforms. This research also touched surveys some of the social media software tools and for completeness introduced social media scraping, data cleaning and sentiment analysis.

Perhaps, the biggest concern is that companies are increasingly restricting access to their data to monetize their content. It is important that researchers have access to computational environments and especially ‘big’ social media data for experimentation. Otherwise, computational social science could become the exclusive domain of major companies, government agencies and a privileged set of academic researchers presiding over private data from which they produce papers that cannot be critiqued or replicated. Arguably what is required are public-domain computational environments and data facilities for quantitative social science, which can be accessed by researchers via a cloud-based facility.

In this research project, the researcher did a detailed review of various existing types of system using text analytics to detect suspicious activity in social media. Basically, the researcher found that the techniques developed in this domain, focuses on text analysis in order to discover the suspicious contents with the deduction of the suspicious behaviour users on the web. Study revealed that we need to analyze the user’s behaviour based on tweets, comments; links shared etc. information available with respect to the user.

Also, suspicious behaviours can be categorized under groups such as terrorism, sexuality, racism, fraud financial laundering, or others. Using this categorization, the corpus of probable suspicious words can be build which will further assist in developing more refined and reliable techniques for detecting these activities. An important aspect of this study divulges that NLP algorithms may prove their calibre and superiority in this domain.

This research presents a research on analysing Facebook and Twitter data in approaches of text mining and social network analysis. With text mining of posts, comments, likes, retweets, tweets, topics and their variations over time have been identiﬁed. Both Facebook and Twitter followers have been analysed and the spread of tweets over Twitter network has been studied. With some initial interesting results identiﬁed, this research will be further studied in future research projects.

**Limitations**

The study encountered the various limitations:

**Operating System:** Due to the huge volumes of data that web forums such as websites and blogs possess there was need of a much faster operating system so as to be able to perform web crawling faster and to go deeper in the web scraping.

**High Data Volumes:** Social Media possess a lot of data and there was need to deploy faster database schemas and mechanism for data storage and processing.

**Programming Language:** Since different programming languages such as Python were used to develop the application, a lot of time was consumed for testing and debugging application related issues.

**Time:** The research required a lot of time in analyzing data and finding the best way to represent the social media mined data.

**Future Work**

The methodology used in this work is general, the tools used are open-source software, and the data used in this work are publicly available on Facebook and Twitter. Therefore, readers can easily replicate the analysis and apply it to Facebook and Twitter accounts that they are interested in. This research can be extended by analysing text from more accounts, analysing social network between them and their followers, and developing an eﬀective methods to ﬁnd relationships between topics/events/posts and variations in time series.

It can also be extended further to investigate how messages spread, estimate their impacts and generate alerts. It would also be interesting to analyse tweets on speciﬁc topics based on Facebook and Twitter hashtags, such as “#ThisFlag”, “#FreeMartha” and “#ZimVotes2018”, and perform sentiment analysis for new legislations and policies. Another possible future work is to study social network with data from multiple social media platforms, such as Twitter, Facebook9 and Google+10, and investigate interactions between government agencies, migration agencies and individuals.

**Summary**

Online social networks have become part of our everyday life and on the average most internet users spend more time on social networks than in any other online activities. Users enjoy using online social networks to interact with other people through the sharing of experiences, pictures, videos, and other types of information. Nevertheless, online social networks have been having a dark side issue with hackers, fraudsters, and online predators, all fond of using online social networks as a platform for procuring their next victim

In this research project, the author has presented scenarios which can be encountered in monitoring suspicious activity by malicious users that affect people’s privacy and well-being both in the virtual world and in the real world. Moreover, the researcher emphasized certain data mining techniques that can be used to monitor suspicious behaviour on social media in particular Facebook and Twitter.

The researcher equally highlighted some past literature on social media monitoring, data mining and applications and mechanisms developed till date. More importantly however, this research successfully brought on board a mechanism that enables monitoring of suspicious behaviour on Facebook and Twitter using data mining.

**Bibliography**

Alami, S and El Beqqali,O (2015). Detecting Suspicious Profiles Using Text Analysis Within Social Media. Journal of Theoretical and Applied Information Technology. Vol.73 No.3

Alami, S and El Beqqali,O (2015). “Cybercrime profiling: Text mining techniques to detect and predict criminal activities in microblog posts.” In 2015 10th International Conference on Intelligent Systems: Theories and Applications (SITA), pp. 1-5. IEEE

Asha & Dr. Balkishan, (2017) A Systematic Review on the Suspicious Profiles Detection on Online Social Media Data. Oriental Journal of Computer Science and Technology. ISSN: 0974-6471, Vol. 10, No. (3) 2017, Pg. 644-652

Adedoyin-Olowe, M., Gaber, M., Stahl, F. (2013): A Methodology for Temporal Analysis of Evolving Concepts in Twitter. In: Proceedings of the 2013 ICAISC, International Conference on Artificial Intelligence and Soft Computing..

Adu M K, Alese B K and Adewale O S. (2014), *Mitigating Cybercrime and Online Social Networks Threats in Nigeria*, Proceedings of the World Congress on Engineering and Computer Science 2014 Vol I WCECS 2014, 22-24 October, 2014, San Francisco, USA

Agarwal A., Xie B., Vovsha I., Rambow O., and Passonneau R., (2011). "Sentiment analysis of twitter data," in Proceedings of the Workshop on Languages in Social Media, Stroudsburg, pp. 30-38.

Aggarwal, N., Liu, H. (2008): Blogosphere: Research Issues, Tools, Applications. ACM SIGKDD Explorations. Vol. 10, issue 1, 20.

Aggarwal, C. (2011): An introduction to social network data analytics. Springer US.

Aiello, L. M., Petkos, G., Martin, C. et al. (2013): Sensing trending topics in Twitter.

Asur, S., and Huberman, B. (2010): "Predicting the future with social network." Web Intelligence and Intelligent Agent Technology (WIIAT), 2010 IEEE/WIC/ACM International Conference on. Vol. 1. IEEE.

Au Yeung, C. M., and Iwata, T. (2011): Strength of social influence in trust networks in product review sites. In Proceedings of the fourth ACM international conference on Web search and data mining (pp. 495-504). ACM.

Bakshy, E., Hofman, J. M., Mason, W. A., Watts, D. J. (2011): Identifying influencers on twitter. In Fourth ACM International Conference on Web Seach and Data Mining (WSDM).

Becker, H., Naaman, M., Gravano, L. (2011): Beyond Trending Topics: Real-World Event Identification on Twitter. ICWSM, 11, 438-441.

Becker, H., Chen, F., Iter, D., Naaman, M., Gravano, L. (2011): Automatic Identification and Presentation of Twitter Content for Planned Events.

Becker, H., Iter, D., Naaman, M., Gravano, L. (2012): Identifying content for planned events across social media sites. In Proceedings of the fifth ACM international conference on Web search and data mining (pp. 533-542). ACM.

Bekkerman, R., McCallum, A. (2005): Disambiguating web appearances of people in a social network. In Proceedings of the 14th international conference on World Wide Web (pp. 463-470). ACM.

Bethard, S., Yu, H., Thornton, A., Hatzivassiloglou, V., Jurafsky. D. (2004): Automatic Extraction of Opinion Propositions and their Holders. In: Proceedings of the AAAI Spring Symposium on Exploring Attitude and Affect in Text.

Boiy, E. and Moens, M. (2009): A Machine Learning Approach to Sentiment Analysis in Multilingual Web Texts. Information Retrieval, 12(5): 526-558.

Bollen, J., Mao, H., Pepe, A. (2011).: Modelling public mood and emotion: Twitter sentiment and socio-economic phenomena. In ICWSM.

Barbier G & Liu H, (2011) Data mining in social media .In C. Aggarwal (ed.), Social Network Data Analytics, DOI 10.1007/978-1-4419-8462-3\_12.

Barbosa L. and Feng J., (2010). "Robust sentiment detection on twitter from biased and noisy data," in Proc. 23rd International Conference on Computational Linguistics, Stroudsburg, pp. 36-44.

Berndtsson, M et al. (2008), *Planning and Implementing your Computing Project - with Success!* Springer

Batrinca B, & Treleaven, P. C (2014), Social media analytics: a survey of techniques, tools and platforms. AI & Soc (2015) 30:89–116

Bolla, Raja Ashok, (2014). "Crime pattern detection using online social media". Masters Theses. 7321. <http://scholarsmine.mst.edu/masters_theses/7321>

Boutet, Antoine, Hyoungshick kim, and Eiko Yoneki. (2012) “What’s in your tweets? I know who you supported in the Uk 2010 general election.” In The International AAAI Conference on Weblogs and Social Media (ICWSM).

Chew C. and Eysenbach G., (2010)."Pandemics in the Age of Twitter: Content Analysis of Tweets during the 2009 H1N1 Outbreak,"

Gladwell, Malcolm, and Clay Shirky. (2011) “From innovation to revolution: Do social media make protests possible?.” Foreign Affairs 90, no. 2 : 153.

Juliet, N.N. (2014). Internet *Freedom in Kenya: Balancing Hate Speech and Free*: Launch of the Internet Freedoms in East Africa 2014, 4(1), 21-25.

Lau, Raymond Yk, and Yunqing Xia. (2013): “Latent Text Mining for Cybercrime Forensics.” International Journal of Future Computer and Communication 2, no. 4 368.

Murugesan, M. Suruthi, R. Pavitha Devi, S. Deepthi, V. Sri Lavanya, and Annie Princy. (2016). “Automated Monitoring Suspicious Discussions on Online Forums Using Data Mining Statistical Corpus Based Approach.” Imperial Journal of Interdisciplinary Research 2, no. 5

Pak A. and Paroubek P.,(2010) "Twitter as a Corpus for Sentiment Analysis and Opinion Mining.," in LREC, vol. 10, Orsay Cedex, 2010, pp. 1320-1326.

Pang B. and Lee L., (2008) "Opinion mining and sentiment analysis," Foundations and trends in information retrieval, vol. 2, no. 1-2, pp. 1-135,

Papadopoulos, S., Kompatsiaris, Y., Vakali, A., Spyridonos, P. (2012): Community detection in social media. Data Mining and Knowledge Discovery, 24(3), 515-554.

Pham, M. C., Cao, Y., Klamma, R., Jarke, M.: A clustering approach for collaborative filtering recommendation using social network analysis. J. UCS, 17(4), 583-604, 2011.

Paul M.J. and Dredze M., (2011) "You are what you Tweet: Analyzing Twitter for public health." 5th International AAAI Conference on Weblogs and Social Media (ICWSM 2011).

Sutton, J. N. (2009). Social media monitoring and the democratic national convention: New tasks and emergent processes. Journal of Homeland Security and Emergency Management, 6(1), 1-20.

Siguenza-Guzman, Lorena, Victor Saquicela, Elina Avila-Ordóñez, Joos Vandewalle, and Dirk Cattrysse. (2015) “Literature review of data mining applications in academic libraries.” The Journal of Academic Librarianship 41, no. 4: 499-510.

Tayal, Deventra Kumar, Arti Jain, Subhi Arora, Surbhi Agarwal, Tushar Gupta, and Nikhil Tyagi. (2015): “Crime detection and criminal identification in India using data mining techniques.” AI & SOCIETY 30, no. 1 117-127.

Twitter. Twitter. [Online]. https://dev.twitter.com/streaming/overview . [cited Oct 20, 2017]

Twitter Inc., support.twitter. [Online]. https://support.twitter.com/groups/50-welcometo-twitter. [cited November 10, 2017]

Twitter, Inc. Twitter. [Online]. https://about.twitter.com/company. [cited November 10, 2017]

Upganlawar, H and Sambhe, N (2017), Surveillance of Suspicious Discussions on Online Forums using Text Data Mining. International Journal of Advances in Electronics and Computer Science, ISSN: 2393-2835 Volume-4, Issue-4.

Valkanas, G., Katakis, I., Gunopulos, D. and Stefanidis, A., 2014, August. Mining twitter data with resource constraints. In Proceedings of the 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)-Volume 01 (pp. 157-164). IEEE Computer Society.

Wafula W G (2016) Social Media Forensics for Hate Speech Opinion Mining. University of Nairobi.

Yang, M., Kiang, M., Ku, Y. (2011). Journal of Homeland Security and Emergency Management. Volume 8, Issue 1

Zhao, J., Liu, K. and Xu, L., (2016). Sentiment Analysis: Mining Opinions, Sentiments, and Emotions Bing Liu (University of Illinois at Chicago) Cambridge University Press, ISBN 9781107017894

Zhou, Y., Reid, E., Qin, J., Chen, H., and Lai, G. (2005). U.S. extremist groups on the Web: Link and content analysis. IEEE Intell. Syst. 20, 5, 44–51