BINDURA UNIVERSITY OF SCIENCE EDUCATION

FACULTY OF SCIENCE AND ENGINEERING

DEPARTMENT OF COMPUTER SCIENCE

CYBER PATROL THREAT DETECTION THROUGH A MULTI-LINGUAL AI ALGORITHM

BY

JUSTICE MUKARO

B1748519

SUPERVISOR: MR ZANO

A RESEARCH PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE BACHELOR OF SCIENCE HONOUURS

DEGREE IN COMPUTER SCIENCE

DECEMBER 2021

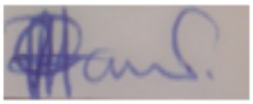
APPROVAL FORM

The undersigned certify that they have supervised the student Justice Mukaro dissertation entitled “CYBER PATROL THREAT DETECTION THROUGH A MULTI-LINGUAL AI ALGORITHM” submitted in Partial fulfilment of the requirements for the Bachelor of Computer Science Honours Degree of Bindura University of Science Education.

16/10/24

……. ……

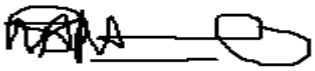
STUDENT SIGNATURE DATE



17/10/24

……... …….

SUPERVISOR SIGNATURE DATE



17/10/24

……. ……..

CHAIRPERSON SIGNATURE DATE

…….. …….. ……/……/……..

EXTERNAL EXAMINER SIGNATURE DATE

ABSTRACT

Billions of tweets are sent out daily in all sectors of life especially in business and governance in multiple languages. Some of the tweets sent out present threats to leaders and also influencers utilise the platform to incite violence and cause civil unrest. This research brings about the creation of a multilingual threat detection algorithm to scour through selected tweets and conduct a threat analysis. Using Twitter APIs, the bot legally extracts tweets and uses the Python language translation function and then it analyses them using Natural Language Processing (NLP). The data is collected, cleaned, analysed and output is given on a word cloud and other statistical representations. The algorithm was created on the Google Collaboratory platform which is a browser-based Python compiler. It uses Google Servers to access RAM, GPU and competent processors. Thus, all the output is automatically stored on a Google Drive. The researcher collected data from the President of the United States, Joe Biden. His tweets were analysed for threats and or abuse of power and also, tweets directed towards were analysed for the same information. The output was presented on a Word Cloud, scatter diagram and a bar graph.

DEDICATION

God’s Work.

ACKNOWLEDGEMENTS

Firstly, all the appreciation goes to the Almighty God who guided me always from the beginning up to my final year and gave me the strength day by day, knowledge and wisdom to pursue my research. My deepest gratitude goes to my supervisor Mr Zano for the continuous and firm support during the research, for his motivation, enthusiasm, immense knowledge and patience. I also appreciate support from BUSE lecturers.

A special thank you goes to my classmates for team work and the enormous effort they put in towards information sharing and for making this journey worthwhile and enjoyable, thank you.

KEYWORDS

- Threat Detection
- Multilingual AI Algorithm
- Twitter bot
- Natural language processing (NLP)
- Python
- Social media (SM)
- Google Collaboratory
- WordCloud
- President of the United States (POTUS)
- Data Mining
- Cyber Patrol

Table of Contents

CHAPTER 1: PROBLEM IDENTIFICATION

1.1 Background

Since the flip into the new century, there has been a widespread adoption of technology in all industries. The cyberspace has expanded rapidly to replace most traditional systems and this has maximised productivity. As we are now in the Fourth Industrial Revolution, Artificial Intelligence, Robotics and Internet of Things (IoT) are taking over as the major technologies of the future (Xu, 2018). The cyberspace is used by multiple high-profile individuals for efficient communication especially on social media. Looking at Twitter in particular, most Head of States, artists and sports people use the platform to engage their constituencies for queries and updates (Curran, 2011).

1.2 Problem Statement

On the flip side of the coin, criminals have also expanded into the same spaces creating new problems. Cases of cyberbullying, racial abuse and fraud are rising at an exponential rate. Sadly, there is a significant rise in mortal threats being delivered to states men and women in different languages, some of which have gone to be lethal resulting in loss of life thereby evidently creating threats National Security. In response, Twitter enacted the violent threats policy which prohibits users from delivering threats on their platform (Twitter, 2019). Shaun Himmerick the producer of the video game, Mortal Kombat ended up deleting his account after threats were delivered on his family (arsTechnica, 2015). The limitation of the Violent Threats Policy is that it is report based thus threat detection is not automatic or live which essentially means the victim has to see the threats first in order for any action to be taken. A strong cybersecurity system is proactive and not reactive therefore there is a gap that needs to be addressed. In addition, there exists a language barrier in the threat detection algorithms that are in existence, most of them only analyse in English thus threats delivered in any other language go undetected.

1.3 Research Questions

1. How efficient are AI algorithms in threat detection during Cyber Patrol?
2. How accurate are the algorithms in threat assessment and evaluation?

1.4 Research Aim

To design and develop a multilingual Cyber Patrol AI algorithm that detects threats to state leaders and automatically create a database of suspicious handles and usernames pending further investigations.

1.5 Research Objectives

1. Create a multilingual threat detection algorithm.
2. Create a database of suspects pending investigation.
3. Evaluation on the accuracy of all threats detected.

By definition, Artificial Intelligence is the theory and development of computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages. As such, an algorithm that detects threats in Spanish and English is to be created to analyse data gathered from Twitter posts.

1.6 Research Methodology

A Quantitative approach will be used in the creation of a data set which encompasses formal and slang keywords that are to be considered as threats. This data is to be gathered via Google Forms.

A Qualitative approach is to be implemented in the statistical analysis of the potential threats detected to check for the accuracy of the algorithm and also its usefulness in stopping imminent threats post investigation.

1.7 Significance of Study

This research stimulates the development of a new algorithm thereby widening the existing body of knowledge. More so, it is a proposed modern day African solution to modern day global challenges.

1.8 Scope of study

The Research will mainly be focusing on twitter comments or opinions for steward bank posts. These comments will be analyzed to help improve customer services and quality of products. The author is working on this project having in mind other business people or organizations that may need this system. The research is only limited to use Twitter APIs and machine learning technology to fulfil this research.

1.9 Ethical consideration of Twitter data

Data extracted from twitter APIs contain personal information which means they are subject to relevant data protection legislation including the UK Data Protection Act (DPA) (ML Williams, 2017). The researcher has to accept that terms of services the terms of services of social media networks provide adequate provision to cover this aspect of the DPA.

1.10 Limitations

- Time needed to carry out the research is limited.
- There many restrictions when using twitter APIs.
- Lack of funding for the research.

1.11 Conclusion

Twitter has risen to be the most prominent social media (SM) platform in the world. A lot of good has been birthed on this platform and an equally disturbing amount of evil is prevailing. This chapter has been on problem identification and it has enlightened several contributions on this research. The next chapter will be the literature review.

CHAPTER 2: LITERATURE REVIEW

2.1 INTRODUCTION

Literature review is the act of assessing and analysing the work of other researchers in a shared research area. (Rhoades, 2011) defines literature review a search and evaluation of the available literature in your given subject or chosen topic area. It documents the state of the art with respect to the subject or topic you are writing about. It synthesises the information in that literature into a summary. The previous chapter was exhibiting the problems in the society emanating from the use of Twitter, and the weaknesses of existing protective measures. Millions of threats are posted daily and detecting them proves to be a daunting task if done manually. This chapter will examine the work that has been done towards automatic threat detection systems and multi-lingual algorithms.

2.2 THREAT DETECTION SYSTEMS AND DATA MINING

Data itself has become a mineral in the digital era, it is rich with useful information that can be utilised in business, politics and of course, in socialising. Data mining is a new discipline lying at the interface of statistics, database technology, pattern recognition, machine learning, and other areas. It is concerned with the secondary analysis of large databases in order to find previously unsuspected relationships which are of interest or value to the database owners (Agha, 2014). Deep Neural Networks (DNN) have been used in the creation of social media threat detection systems. In the context of increasing cyber threats and attacks, monitoring and analysing network security incidents in a timely and effective way is the key to ensuring network infrastructure security, thus the system focuses on infrastructural security and not human security (Fang, 2020).

2.3 TWITTER VIOLENT THREAT POLICY

Healthy conversation is only possible when people feel safe from abuse and don't resort to using violent language. For this reason, we have a policy against threatening violence on Twitter. We define violent threats as statements of an intent to kill or inflict serious physical harm on a specific person or group of people. Under this policy, you cannot state an intention to inflict violence on a specific person or group of people. We define intent to include statements like "I will", "I'm going to", or "I plan to", as well as conditional statements like "If you do X, I will". Violations of this policy include, but are not limited to threatening to kill someone, threatening to sexually assault someone, threatening to seriously hurt someone and/or commit any other violent act that could lead to someone's death or serious physical injury and asking for or offering a financial reward in exchange for inflicting violence on a specific person or group of people (Twitter, 2019).

2.4 MULTILINGUAL ARTIFICIAL INTELLIGENCE ALGORITHMS

Over the past two decades, a significant number of multilingual algorithms have been created in Natural Language Processing (NLP). Among them is the Bilingual Evaluation Understudy (BLEU), is a score for comparing a candidate translation of text to one or more reference translations. Although developed for translation, it can be used to evaluate text generated for a suite of natural language processing tasks (Papineni, 2002). The world's leading search engine, Google created an efficient translation algorithm which covers most of the popular natural languages spoken around the globe. Google Translate uses a Neural Machine Translation (GNMT**)** algorithm which uses a large artificial neural network capable of deep learning. By using millions of examples, GNMT improves the quality of translation, using broader context to deduce the most relevant translation. The result is then rearranged and adapted to approach grammatically based human language (Ghasemi, 2016).

2.5 MACHINE LEARNING

(Emerj Artificial Intelligence, n.d.), defines Machine Learning as the science of getting computers to learn and act like humans do, and improve their learning over time in autonomous fashion, by feeding them data and information in the form of observations and real-world interactions. Furthermore, Machine learning algorithms create a computational template centred on test data, so that projections or choices can be made without explicit programming for the assignment. It is subset of artificial intelligence based on the concept that, with minimal human interference, systems can learn from information, identify trends, and make choices. Through machine learning a software program can understand its surroundings and make choices appropriately on the basis of what they obtain.

2.5.1 Types of Machine Learning

The three most common kinds of machines that are monitored, unattended and strengthening learning formally known as supervised, unsupervised and reinforcement learning.

2.5.2 Supervised Machine Learning

This algorithm consists of a target or outcome variable (or dependent variable) which is to be predicted from a given set of predictors (independent variables). Using these set of variables, a function is generated that map inputs to desired outputs. The training process continues until the model achieves a desired level of accuracy on the training data. Examples of Supervised Learning: Regression, Decision Tree, Random Forest, KNN, Logistic Regression etc. (Abdi, 2016)

2.5.3 Unsupervised Learning

Is a term that refers to Hebbian teaching, allied with teacher-free learning, also known as self-organization, is a method of modelling input probability density (Siadati, 2018). A central framework of unmonitored learning is statistical density estimation, although unsupervised teaching involves many other fields involving the summary and explanation of data characteristics.

2.5.4 Reinforcement Learning

Is a machine learning zone involved with how software officials should act in an area to maximize some cumulative compensation concept? It varies from supervised teaching in that marked input / output duos do not need to be present and sub-optimal activities do not need to be clearly fixed. The focus is instead on discovering equilibrium between exploring and exploiting present understanding (Du, 2014).

2.6 REVIEW OF SIMILAR WORK

Under this section, applications that perform sentiment analysis for social networks will be discussed and analysed by the author.

Open Dover is a web service that performs sentiment analysis tagging on any piece of text provided. The user provides the content that he or she wants to be analysed. The content is returned as feedback with sentiment tags and negative or positive appraisal of the whole file. In order to analyse, Open Dover make use of a knowledge base of opinion words, domain related words and intensifiers and distinguishes between context dependent and context independent opinion words. In addition, Open Dover tries to identify some commonly used domains and evaluates a piece text taking into consideration these domains. Unfortunately, the result is not always as expected, since there is a limit in the available domains ready to be used.

(Hemalatha, 2012) shows an approach for pre-processing Twitter row data following simple steps and demonstrates how to prepare the data for training using machine learning. This approach eliminates unnecessary data or text such as slang, abbreviation, URL, special characters from the linguistic data and also reduces the size of data set which is fast to analyse.

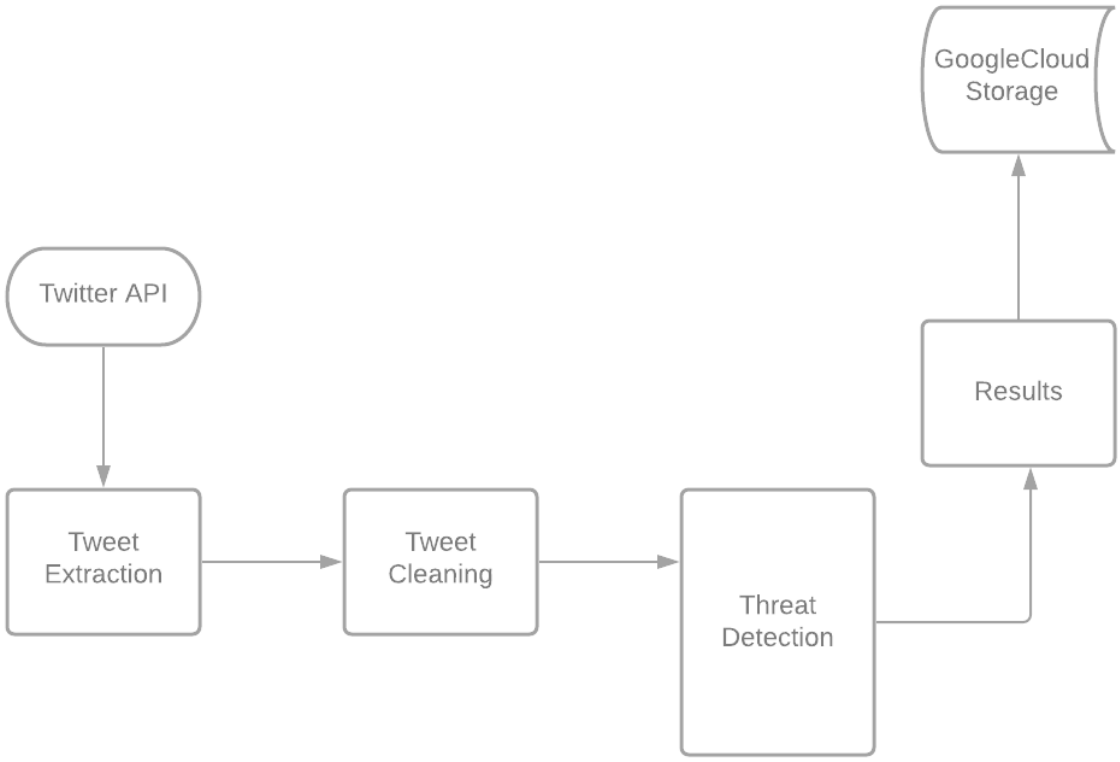
2.7 Review of Existing System

There is no single machine learning model or technique that is appropriate to every opinion mining problem. The same holds also for sub problems like mining Twitter data. These systems are also expensive to purchase. They also analyse document as a whole and give overall sentiment analysis.

2.8 Proposed System

An intelligent device which draws data from Twitter through an Application Programming Interface (APIs) following a search query as per the Twitter Community guidelines. The tweets are pre-processed or cleaned removing special characters, URLs and other unnecessary characters. Using natural language processing, the system will carry out threat assessment of pre-processed data classifying it into Lethal or Non-Lethal and create a database of all suspicious accounts prior to further investigations. Statistical representation of the results is shown on the output box.

2.9 Conceptual Model



2.10 Conclusion

This chapter gave a review of different researches undertaken by different authors to address the issue of multilingual analysis and threat detection on twitter data, in a very short space of time and with very low cost. The researcher did a full assessment of the algorithms used by the existing systems and found some gaps which he then filled and improves using Twitter APIs and Machine learning technique. The next chapter is going to look at research methodology, which is how the research was executed.

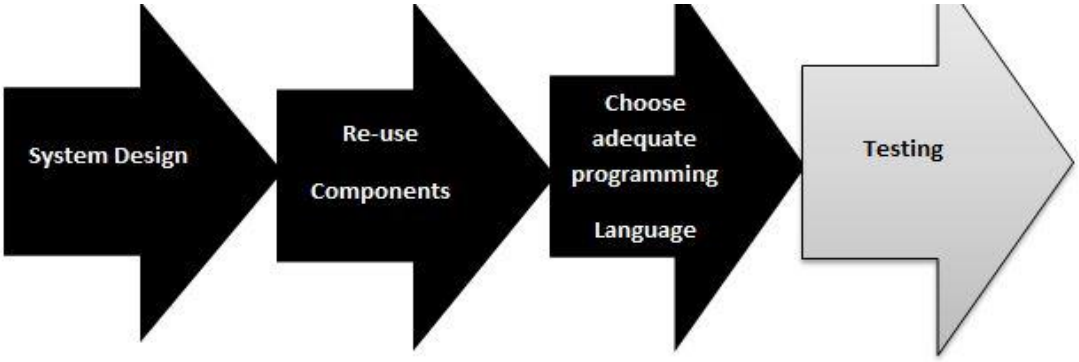
CHAPTER 3 RESEARCH METHODOLOGY

3.1 Introduction

This chapter defines the strategies and tools used to achieve the research objectives. With the help of the information attained in the previous chapter of literature review, the researcher formulated the necessary methods to build a solution and be able to make choices among competing strategies to achieve the expected results of the research.

3.2 Research Design

Research design refers to the framework of market research methods and techniques that are chosen by a researcher. The design that is chosen by the researchers allow them to utilise the methods that are suitable for the study and to set up their studies successfully in the future as well (Akhtar, 2016). The core objective of this stage is to ensure that an operative, proficient, sustainable and reliable system is designed. The researcher used build research methodology which consists of building an artifact either a physical artifact or a software system to demonstrate that it is possible. To be considered research, the construction of the artifact must be new or it must include new features that have not been demonstrated before in other artifacts (Amaral, 2020). The methodology consists of four phases to develop the systems, which are illustrated below:



System Design – After gathering system requirements and well understood them, the researcher designed the system and considered a modular approach, which allows for simplification of system testing so as to increase flexibility and re-use potential of the system.

Re-use Components – The researcher adopted already available components that are free for instance in this research use of libraries for the different modules that are already available from previous system related to the proposed system. This saves system development time.

Choose Adequate Programming Language – involves use of a programming language which is more adequate for the building of a specific system, the researcher chose Python. Important factors to consider when choosing a programming language to work with include: required run-time speed (compiled vs. interpreted languages), expressiveness (imperative vs. functional vs. declarative languages), reliability (e.g., run-time checks, garbage collection) and available libraries. In consideration of the above, the researcher utilised Google Colab to access competent resources (RAM, CPU, GPU) online for free offered by Google.

Testing – it involves testing the developed system each time and the methodology requires testing modules first. This way future changes can be tested when they are introduced into the system.

3.3Targeted Population

According to (Schutt, 2012) target population refers to a set of components or population that was sampled and to which the researcher would like to take a broad view to any study findings. The target population for this research were three twitter accounts, President of the USA (Joe Biden), researcher twitter account (Search API) and one testing twitter account.

3.4 Sample and Sampling Techniques

The research made use of convenience sampling technique to select one influential individual (Joe Biden) for this research to be successful. Convenience sampling is selecting participants because they are often readily and easily available (Revoli, 2020). This was due to limited time and resources hence cost and time effective for this research.

3.5 Data Collection

Data collection is the process of gathering and measuring information on variables of interest, in an established systematic fashion that enables one to answer stated research questions, test hypotheses, and evaluate outcomes (Kabir, 2020). The researcher was able to collect the required data using interviews, observations and background reading of the actual twitter posts and comments, observing how the system behaves on a daily basis.

3.6 Observation

Observation is a fundamental way of finding out about the world around us (Kabir, 2020). It is a systematic data collection approach in research. The Researcher used all of his observation ability to examine people comments or reactions in natural settings or naturally occurring situations.

Advantages

- They provide ecological valid recordings of natural behaviour or reactions.
- Give broader overview of the situation.

Disadvantages

- The use of observation was time intensive by nature and this way was very costly since I was using Twitter social media platform. The observations are subject to bias if they are not successfully implemented or documented.

3.7 Background Reading

For understanding exactly how other researchers dealt with identical problem(s), the author carried out a thorough and detailed background reading. A gap still exists in the already existing solutions and researches that were performed. The documents of currently existing systems, online journals, eBooks and the internet were the key sources the researcher used for background reading. Due to movement restrictions because of Covid-19 pandemic, this method becomes the easiest one to gather information for this research to be successful.

Advantages of Background Reading

- An enriched or better understanding of the problem in question was extended helping in covering existing gap.
- Provides the researcher with an insight into the topic and understanding of the gap in existing system.

Disadvantages of Background Reading

- Some documents are hard to acquire due to payment requirements.
- Documents can contain irrelevant information which does not focus much on the research.

3.8 Requirements Analysis

Requirement analysis is critical to the failure or success of a project and the formulated requirements need to be realistic, documented, testable, actionable, traceable, and measurable, related to identified business needs and detailed enough to facilitate the system design. Since it defines the expectations of the users for a system that is to be built, there is great need for system developers to involve the potential users of the system so that they can keep pace with dynamic changing needs of users. It is essential at this point to record all the functional and non-functional specifications of the required system.

3.9 Functional Requirements

Functional requirements are the processes and functionality that the new framework or system must meet. Functional requirements are clear depiction or indication of the considerable number of service and activities that the new system should give or meet. Along these lines basing on the documentation given the proposed framework has some functional requirements as below:

- The system should be able to extract tweets.
- The system should be able to do threat analysis there by classifying them as Threatening or Non-Threatening.
- The system should be able to exhibit the mined tweets on a Word Cloud.

3.10 Non-functional Requirements

These are the standards that a framework or research is expected to meet without putting any attention on the project development process. Non-functional requirements are often referred to as quality requirements and used to judge the performance of a system rather than its intended behaviour. The non-functional requirements of the proposed system are as follows:

Maintainability

The new framework must be effectively adjusted and equipped for giving new abilities and easy to maintain.

Reliability

Handle framework failures when in operation.

Error handling

The system should rapidly recognise and troubleshoot every error when they occur during operation.

Hardware Requirements

The following are the hardware components that are required in developing the proposed System;

- At least a core i3 processor.
- Windows 8 and above capable computer with reliable network driver.

Software Requirements

1) Windows 8 or better operating system
2) Python development tools

3.11 System Development

This describes the overview of the system or research and how it was developed so as to produce desired results. Thus, it specifies all the software tools and models used in the development of the system.

3.12 Development Methodology

The researcher used Prototyping model which involves building software application prototypes which displays the functionality of the system under development, but may not actually hold the exact logic of the original software.

The basic stages when developing a prototype involves Basic Requirement Identification, Developing the initial Prototype, Review of the Prototype, Revise and Enhance the Prototype.

Basic Requirement Identification:

At this stage, the author understood the very basics system requirements of the system especially in terms of user interface.

Developing the initial Prototype:

The researcher developed the initial prototype where the basic requirements are showcased and user interfaces are provided. User at this stage was able to input search queries into the system as well as number of tweets to be mined.

Review of the Prototype:

The researcher presented the prototype to the BUSE Innovation Hub. Feedback was collected in an organized manner and used for further enhancements of the system under development.

Revise and Enhance the Prototype:

The feedback and the review comments are discussed at this stage and some negotiations are carried out with the customer based on factors like time and budget constraints and technical feasibility of the actual implementation. The changes accepted are again incorporated in the new Prototype developed and the cycle repeats until the customer expectations are met.

3.13 Prototype Model

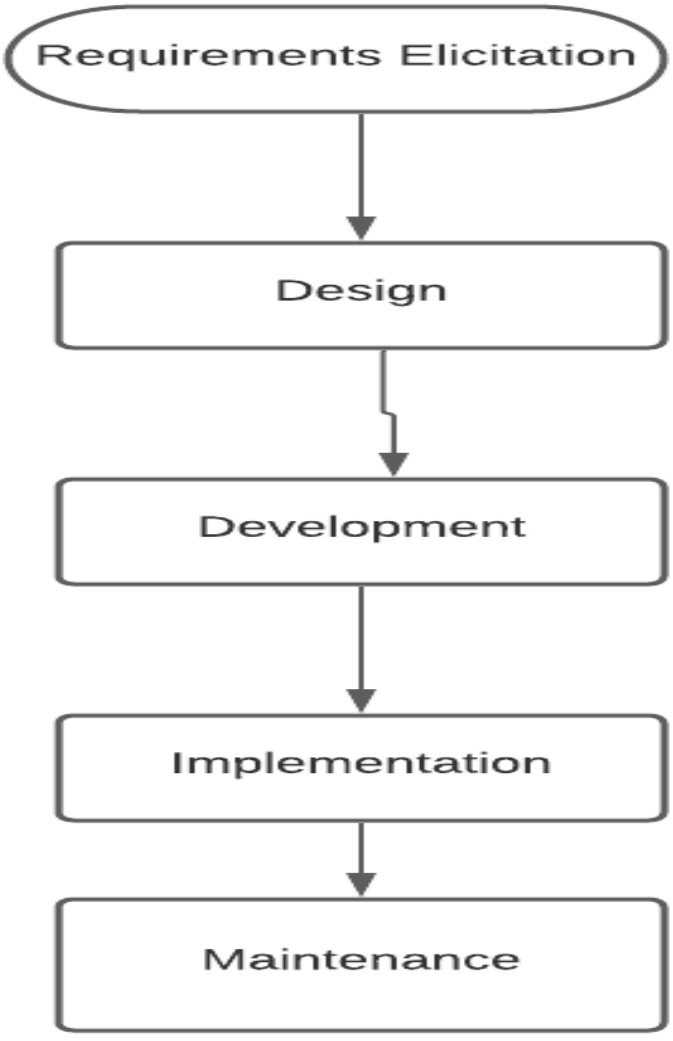


Figure 1

3.14 CyberPatrol Flowchart

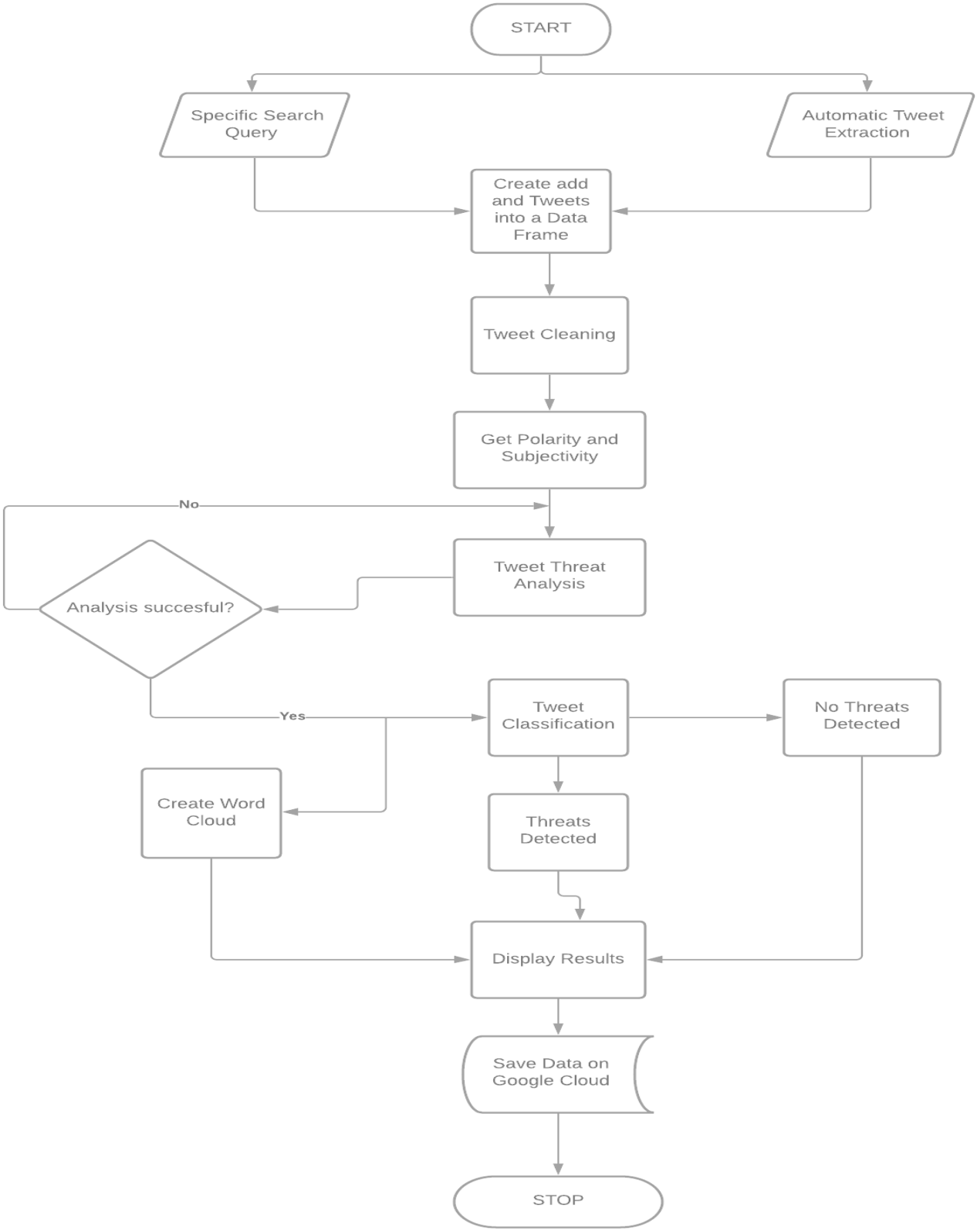


Figure 2

3.15 Component Diagram

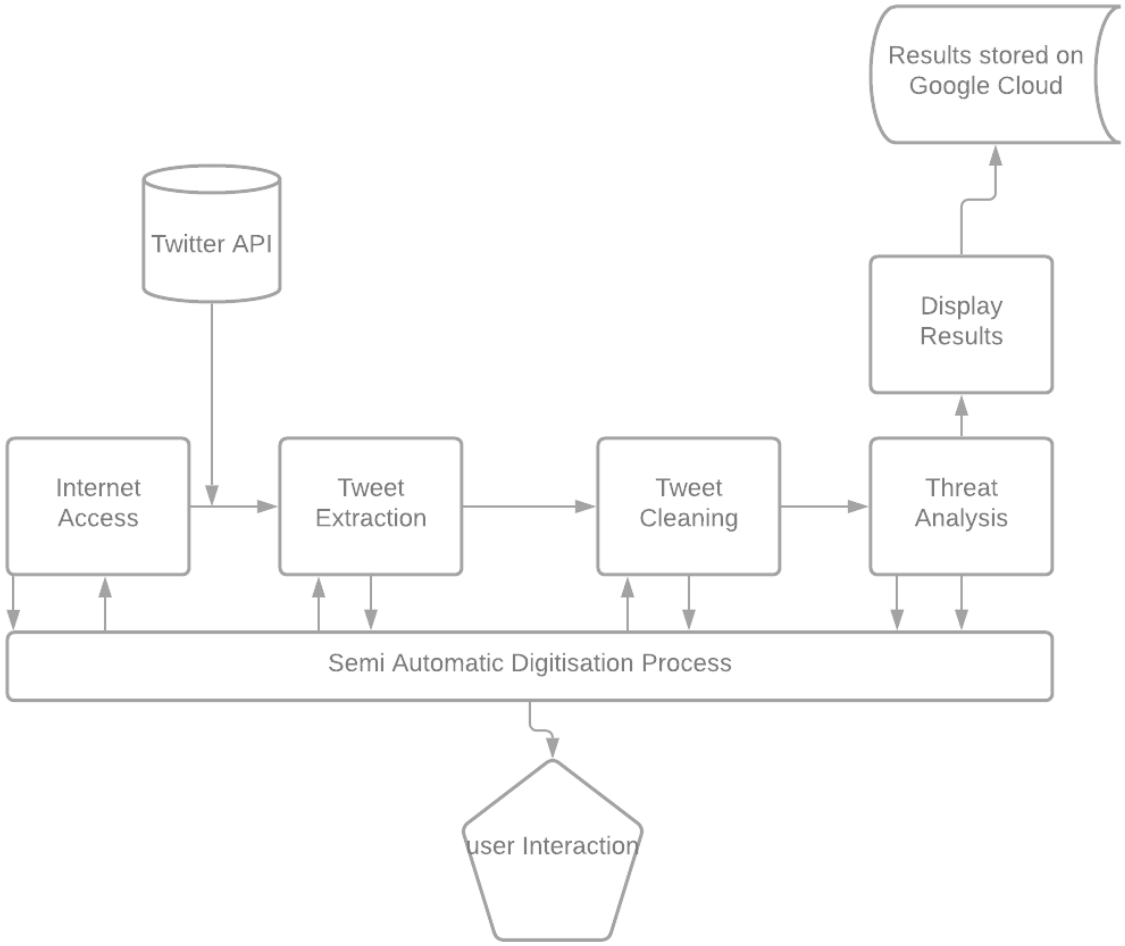


Figure 3

3.16 Summary on how the system works

The component diagram illustrates how the system works. There exists an automatic bot which has been granted access credentials (Secret Keys). The bot legally extracts tweets (raw data) from Twitter. The data gathered is saved into a Data Frame. The tweets in the data frame are then cleaned off usernames and handles as per Twitter guidelines, this is done through Natural Language Processing and Regular Expressions. The data is then analysed for threats through NLP basing on the Lexicon database. The classification of data as either Threatening or Non-Threatening is done and the results are displayed on a Word Cloud. Furthermore, the processed data is expressed in statistical diagrams (Scatter and Bar charts). The final output is saved on Google Cloud.

3.17 Conclusion

This chapter mainly focused on the methodology used in the development of the system and how it was designed as well as implemented. Different techniques were used to come up with the system.

CHAPTER 4: RESULTS ANALYSIS

4.1 Introduction

The need to analyse the efficiency of the developed solution arose when author had successfully implemented the system. Accuracy, performance and reliability used to determine the efficiency and effectiveness of the developed solution. The developed solution's behaviour was also well observed by the researcher under the different times and the outcome was recorded.

4.1 Testing

Testing is of paramount importance in the development process and this chapter shows the tests that were undertaken and the result they produced. The testing is thus measured against the functional and non-functional requirements of this research as outline in the previous chapter.

4.21 Black box and White Box testing

Black box testing enables a user without the knowledge of the internal structure of the system to test it against the functional and sometimes the non-functional requirements of the system. The research mainly focused on extracting tweets made by a high-profile individual and also tweets directed to them and perform a threat assessment on whether the tweets are threatening to either the individual or harmful to the public. Thus, the main purpose of black box testing was to test if the system worked as per expected in requirement section. White box testing is the software testing method in which internal structure is being known to tester who is going to test the software.

4.3 Presentation of Results

4.3.1 A snippet of the algorithm importing the necessary libraries.

```
[1]  #Importing Required Libraries

     import tweepy
     from textblob import TextBlob
     from wordcloud import WordCloud
     import pandas as pd
     import numpy as np
     import re
     import matplotlib.pyplot as plt
     plt.style.use('fivethirtyeight')
```

Figure 4, Importing Libraries

4.3.2 Extraction of POTUS Tweets

After importing libraries and successfully authenticating into twitter. The algorithm goes on to extract a 100 tweets from President of the United States (POTUS), Joe Biden's twitter account. The tweets are to be extracted and translated into English language through the Lang function. The algorithm goes on to print the five most recent tweets.

```
#Extract Tweets
posts = api.user_timeline(screen_name = 'JoeBiden', count = 100, lang ='en', tweet_mode ='extended')

# Print five recent tweets
i=1
print ("Show five recent tweets: \n")
for tweet in posts [0:5]:
 print( str (i) + ')' + tweet.full_text + '\n')
 i=i+1
```

```
Show five recent tweets:

1)As we move into the winter and face the challenges of a new variant, this is a moment for us to put the divisiveness

2)Everybody talks about the price tag for the Build Back Better Act.

But here's the truth: If you make less than $400,000 a year, your taxes will not go up one single penny under our plan

3)Health care should be a right in this country, not a partisan issue. Diabetes, cancer, and other diseases don't care

This is about being able to afford the prescription drugs you need. Congress must come together and finish the job.

4)RT @POTUS: Today, I was briefed by FEMA on our response to the tornadoes and extreme weather that impacted multiple

5)The pandemic has put a strain on our global supply chains and caused prices to rise around the world.
```

Figure 5, Extracting Tweets from the POTUS Account

4.3.3 Extracting Tweets on the POTUS Query

The next step is for the algorithm to successfully extract tweets directed towards POTUS through the Twitter search API.

```
[12] #get tweets about a specific topic

     public_tweets = api.search('JoeBiden')
     # Print 6 recent tweets
     i=1
     print ("Show six recent tweets: \n")
     for tweet in public_tweets [0:6]:
      print( str (i) + ')' + tweet.text + '\n')
      i=i+1
```

```
Show six recent tweets:

1)@JoeBiden Release a statement on how the New England Patriots dynasty is back, John, or else I'll go back in time and vote Green Party

2)RT @Yohannes_v: @VOA_Wandera @CNN 🚨ATTENTION: WAR CRIMES!!

#TIGRAYGENOCIDE!!

@SecBlinken @USAmbUN @UN @antonioguterres @JustinTrudeau @Sc…

3)https://t.co/JfneyBgspg https://t.co/7ZskyXEalC

4)RT @LoveUganda7: @Louise94622028 @Musoke78548359 @chidi_blyden @usmissionuganda @UgandaMediaCent @DeptofDefense @KagutaMuseveni @dw_africa1…

5)RT @ScottMadin: hey there Mr. President, it's been like a year and nine months, just circling back to check where we're at on this

6)@JoeBiden Sat at home for 1.5 years, wore mask, 3x vaxxed along with entire family, contributed generously to you &amp;amp;… https://t.co/VeD2OCemS5
```

Figure 6, Extracting Tweets directed to POTUS

4.3.4 Creation of DataFrame

The algorithm successfully extracted tweets about POTUS and it displayed all six of them in English. All the tweets were compiled and stored in a DataFrame.

```
]  # Creating a dataframe
   df = pd.DataFrame([tweet.full_text for tweet in posts], columns= [ 'Tweets'])
   #df = pd.DataFrame([tweet.text for tweet in public_tweets], columns= [ 'Public'])

   #Show first five rows of data
   df.head()
```

	Tweets
0	As we move into the winter and face the challe...
1	Everybody talks about the price tag for the Bu...
2	Health care should be a right in this country,...
3	RT @POTUS: Today, I was briefed by FEMA on our...
4	The pandemic has put a strain on our global su...

Figure 7, DataFrame

4.3.6 Subjectivity and Polarity

With all the tweets now in the DataFrame, the data is then cleaned off mentions and other related hyperlinks as per Twitter guidelines and also to prepare it for Natural Language Processing. Basing on the definition of words in the Lexicon Database, the tweets are allocated a degree of polarity and subjectivity as illustrated below,

```
#Create a function to get polarity
def getPolarity(text):
  return TextBlob(text).sentiment.polarity

#Create two new columns
df['Subjectivity'] = df ['Tweets'].apply(getSubjectivity)
df['Polarity'] = df ['Tweets'].apply(getPolarity)

# show the new dataframe and its columns
df
```

	Tweets	Subjectivity	Polarity
0	As we move into the winter and face the challe...	0.463636	-0.090909
1	Everybody talks about the price tag for the Bu...	0.195238	0.065476
2	Health care should be a right in this country,...	0.511905	0.220238
3	RT @POTUS: Today, I was briefed by FEMA on our...	0.500000	-0.062500
4	The pandemic has put a strain on our global su...	0.125000	0.125000
...	...	...	...

Figure 8. Subjectivity and Polarity

4.3.7 Word Cloud

To enable human analysis of the tweets, a Word Cloud is created which plots the words with the highest frequency in a relatively larger text on an image and vice versa.

```
[ ]  #Plot The Word Cloud
     allWords = ' '.join( [twts for twts in df ['Tweets']])
     wordCloud = WordCloud (width =600, height=400, random_state= 21, max_font_size= 119).generate(allWords)

     plt.imshow(wordCloud, interpolation = 'bilinear')
     plt.axis('Off')
     plt.show
```

```
<function matplotlib.pyplot.show>
```

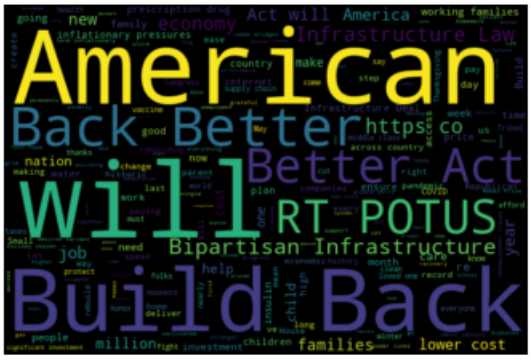


Figure 9, Word Cloud

4.3.8 Threat Analysis

Moving on, the threat analysis function is invoked, the tweets are scanned for threats and results displayed.

```
# invoke analysis function
df['Analysis'] = df['Polarity'].apply(getAnalysis)

#show dataframe
df
```

	Tweets	Subjectivity	Polarity	Analysis
0	As we move into the winter and face the challe...	0.463636	-0.090909	Threats Detected
1	Everybody talks about the price tag for the Bu...	0.195238	0.065476	No threats Detected
2	Health care should be a right in this country,...	0.511905	0.220238	No threats Detected
3	RT @POTUS: Today, I was briefed by FEMA on our...	0.500000	-0.062500	Threats Detected
4	The pandemic has put a strain on our global su...	0.125000	0.125000	No threats Detected

Figure 10, Threat Analysis

4.3.9 Print all Threats

The algorithm then invokes a function to print all tweets which detected as potentially threatening or alarming.

```
[ ] #Print all  Threats
    j=1
    sortedDF= df.sort_values(by = ['Polarity'], ascending = 'False')
    for i in range(0, sortedDF.shape[0]):
      if sortedDF['Analysis'][i] == 'Threats Detected':
        print( str (j) + ' )' + sortedDF['Tweets'][i])
        print ()
        j = j+1
```

```
1 )As we move into the winter and face the challenges of a new variant, this is a moment for us to put the divisiveness behind us. This is a moment for the natior

2 )RT @POTUS: Today, I was briefed by FEMA on our response to the tornadoes and extreme weather that impacted multiple states this weekend. We…

3 )RT @POTUS: This morning, I was briefed on the devastating tornadoes across the central U.S. To lose a loved one in a storm like this is an…
```

Figure 11, List of Threats Detected

4.3.10 Print all Non-Threats

The algorithm then invokes a function to print all non-threatening tweets.

```
[ ] #Print all Non Threats
    j=1
    sortedDF= df.sort_values(by = ['Polarity'])
    for i in range(0, sortedDF.shape[0]):
      if sortedDF['Analysis'][i] == 'No threats Detected':
        print( str (j) + ' )' + sortedDF['Tweets'][i])
        print ()
        j = j+1
```

```
The Bipartisan Infrastructure Law and Build Back Better Act will help lower costs and build an economy that works for everyone.

59 )It was good to be back in Detroit. https://t.co/DXKpvHRteJ

60 )Together, the Bipartisan Infrastructure Law and Build Back Better Act will create millions of jobs, lower costs for working families, and meet th

These are transformative investments in the American people and our future.

61 )Tribal lands have been chronically underfunded for generations. I’m proud to say that the Bipartisan Infrastructure Law will be the single larges
```

Figure 12, List of Non-Threatening Tweets

4.3.11 Scatter Diagram

A list of all the tweets that are considered to be non-threatening tweets is successfully printed. A scatter diagram of the Threat Analysis and Polarity is successfully plotted.

```
<function matplotlib.pyplot.show>
```

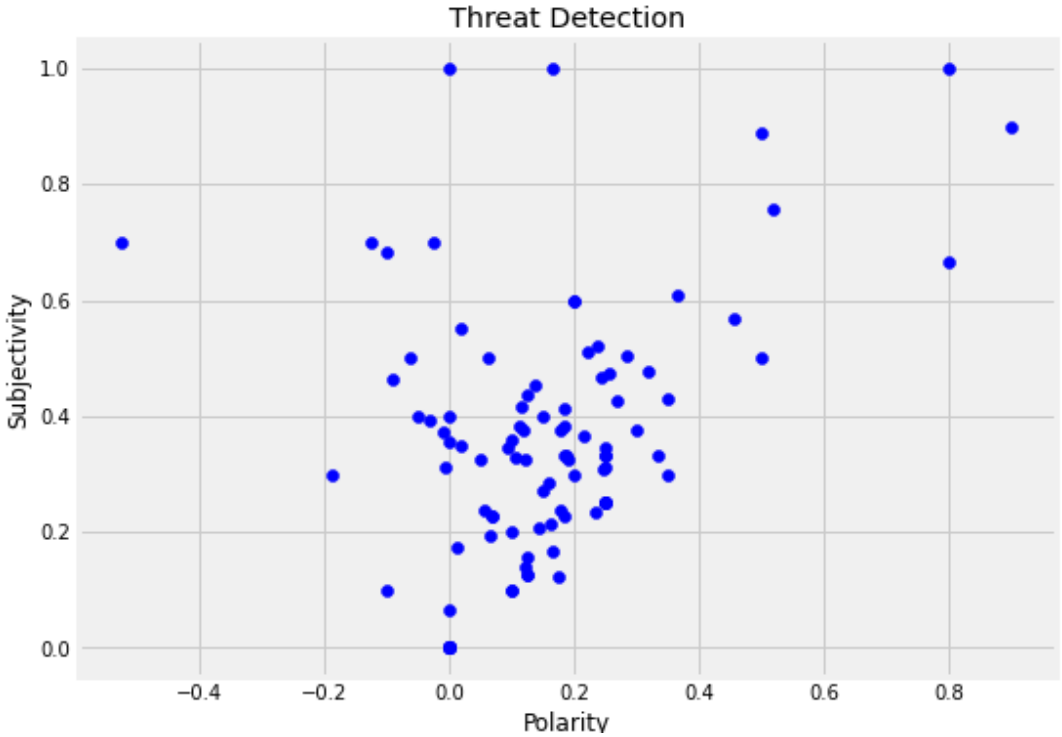


Figure 13, Threat Detection-Polarity

4.3.12 Evaluation of Results

As per the objectives of the research, the algorithm evaluated its efficiency and accuracy as illustrated below

```
[ ]  # Get the percentage of Non Threats
     ptweets= df[df.Analysis == 'No threats Detected']
     ptweets = ptweets['Tweets']

     round (ptweets.shape[0] / df.shape[0]*100, 1 ) # Rounded off to 1dp
```

```
     75.0
```

```
[ ]  # Get the percentage of  Threats
     ntweets= df[df.Analysis == 'Threats Detected']
     ntweets = ntweets['Tweets']

     round (ntweets.shape[0] / df.shape[0]*100, 1 ) # Rounded off to 1dp
```

```
     12.0
```

Figure 14, Evaluation of Results

4.3.13 Value Counts

The algorithm successfully detected seventy-five percent of the tweets as non-threatening, twelve percent were detected as threatening and alarming and the remaining thirteen percent was considered to be neutral. A statistical bar chat of the evaluation was plotted, see below;

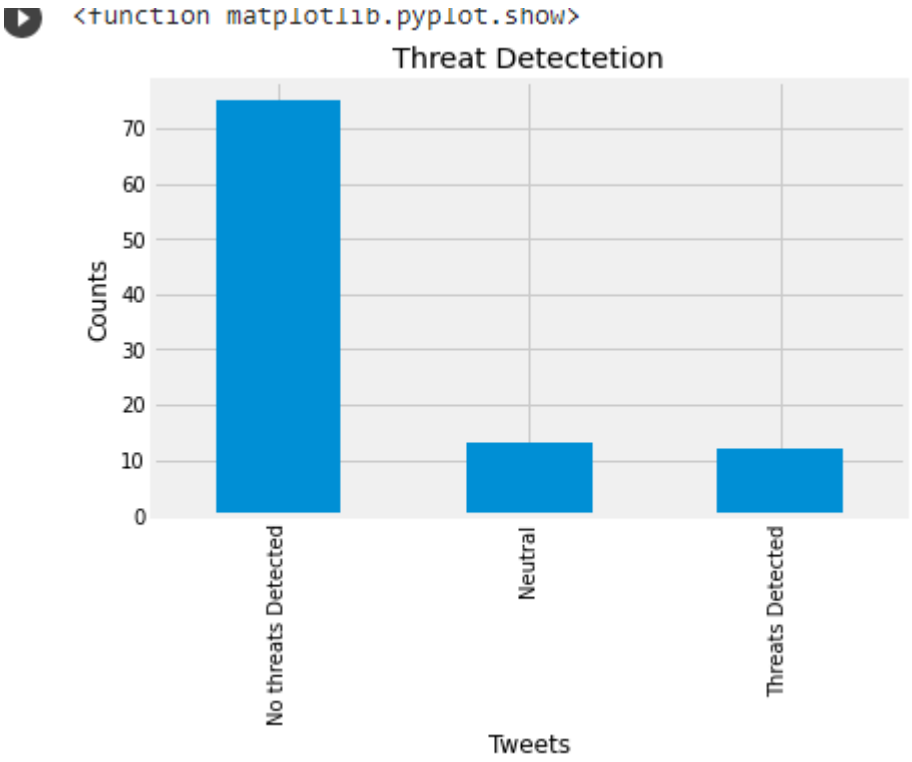


Figure 15, Value Count

The results shown above exhibit that the algorithm was able to extract tweets and automatically translate all of them into English through the python Lang function. The tweets were efficiently cleaned using regular expressions only to leave the natural language required for Natural Language Processing (NLP). The tweets were efficiently assessed by the threat analysis function. The data was presented on a word cloud and statistical graphs, and an evaluation on the accuracy and efficiency of the system was conducted.

4.4 Conclusion

From this research, it can be concluded that Multilingual AI algorithms are efficient in threat Detection systems and they can tremendously contribute to the security of both high-profile individuals and the state (citizens) as well. The algorithm effectively met all the functionalities and goals of the research and accurate results were attained.

CHAPTER 5: CONCLUSIONS AND RECOMMENDATIONS

5.1 Introduction

The previous chapter focused on presentation, evaluation and analysis of the results attained from the research. This chapter focuses on recommendations, conclusions and future work as far as threat detection and assessment is concerned. In addition, this chapter looks at the drawbacks faced by the researcher during implementing the research system under consideration.

5.2 Aims and Objectives Realisation

The aim of the research was to create a multilingual AI algorithm that assists in threat detection, case of state leaders. The objectives were to create the algorithm, automatically document suspicious usernames and evaluate the results. The creation of the algorithm was achieved through the use of the Python Lang function which drew translations from Google Translate. The automatic logging of suspicious handles into a database was not achieved as it is prohibited by Twitter Community guidelines particularly on user privacy. More so, during the creation of the Twitter Developer Account creation, the researcher's account was registered as a student/scholar which meant a limitation to the amount of data that could be extracted. The algorithm performed the evaluation as intended by the goals and objectives.

5.3 Future Work

The future of online threat detection lies in solutions that are closer to home. The Lexicon database used in the analysis of data in this research, largely supports European Languages (English, Spanish). The enhancement of this algorithm to support African languages will contribute tremendously to its efficiency.

5.4 Challenges Faced

The researcher faced a challenge in tweet extraction, each query and command had to be manually activated or invoked. Furthermore, the Twitter Developer account was registered as a scholar account, this meant multiple restrictions on the number of tweets that can be extracted and even more especially on their intended use. The Twitter privacy policy prohibits the documentation of their users account names and or handles thus the creation of a database was unsuccessful. The COVID-19 pandemic presented a myriad of problems with the researcher falling ill and stiff economic situations meant scarce resources.

5.5 Recommendations

It is recommended that researchers work closely with Innovation Hubs such that the projects are enhanced and taken to industry. In regards to the algorithm, a powerful server with adequate resources is required to efficiently execute all the instructions. More so, there should be a reliable and an un-interrupted internet supply because the whole system requires stable internet to operate efficiently and effectively.

5.6 Conclusion

The researcher successfully created the multilingual intelligent algorithm which effectively detect threats from extracted tweets. The researcher grew as an innovator and managed to network. In a nutshell, this research was a success.

References

Abdi, A. (. (2016). Three types of Machine Learning Algorithms.

Agha, S. &. (2014). An Introduction to Data Mining Technique. *IJAETMAS. 3. 5.* .

Akhtar, I. (2016). Research Design. .

Amaral, I. (2020).

arsTechnica. (2015). *arsTechnica*. Retrieved from https://arstechnica.com: https://arstechnica.com/gaming/2015/03/mortal-kombat-producer-quits-twitter-over-violent-threats-to-family/

Curran, K. &. (2011). The Role of Twitter in the World of Business. . *IJBDCN. 7. 1-15. 10.4018/jbdcn.2011070101.* .

Du, K.-L. &. (2014). Reinforcement Learning. . *10.1007/978-1-4471-5571-3_18.*

Emerj Artificial Intelligence. (n.d.). *Emerj Artificial Intelligence Research.* . Retrieved from https://emerj.com: https://emerj.com/ai-glossary-terms/what-is-machine-learning/

Fang, Y. &. (2020). Detecting Cyber Threat Event from Twitter Using IDCNN and BiLSTM. . *Applied Sciences. 10. 5922. 10.3390/app10175922.*

Ghasemi, H. &. (2016). A Comparative Study of Google Translate Translations: An Error Analysis of English-to-Persian and Persian-to-English Translations. . *English Language Teaching. 9. 13. 10.5539/elt.v9n3p13.*

Hemalatha, I. D. (2012). "Case Study on Online Reviews Sentiment Analysis Using Machine Learning Algorithms.". *International Journal of Innovative Research in Computer and Communication Engineering 2.2 (2014):3182-3188.*

Kabir, L. (2020).

Papineni, K. &. (2002). BLEU: a Method for Automatic Evaluation of Machine Translation. . *10.3115/1073083.1073135.*

Revoli, N. (2020). Sampling Techniques Simplified.

Rhoades, E. &. (2011). Literature Reviews. *The Volta Review. 111. 354-369. 10.17955/tvr.111.1.677.* .

Schutt, J. (2012). Target Population.

Siadati, S. (2018). What is Unsupervised Learning. *10.13140/RG.2.2.33325.10720.*

Twitter. (2019). *Twitter.* Retrieved from Www.Twitter.com: https://help.twitter.com/en/rules-and-policies/violent-threats-glorification

Xu, M. &. (2018). The Fourth Industrial Revolution: Opportunities and Challenges. . *International Journal of Financial Research. 9. 90. 10.5430/ijfr.v9n2p90.*