DETECTING ARABIC TERRORISM MESSAGES ON TWITTER USING MACHINE LEARNING

ALHARBI, NORAH MUTEB S

FACULTY OF COMPUTER SCIENCE AND
INFORMATION TECHNOLOGY
UNIVERSITY OF MALAYA
KUALA LUMPUR

2019

DETECTING ARABIC TERRORISM MESSAGES ON TWITTER USING MACHINE LEARNING

ALHARBI, NORAH MUTEB S

DISSERTATION SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF COMPUTER SCIENCE [APPLIED COMPUTING]

FACULTY OF COMPUTER SCIENCE AND
INFORMATION TECHNOLOGY
UNIVERSITY OF MALAYA
KUALA LUMPUR

2019

UNIVERSITY OF MALAYA
ORIGINAL LITERARY WORK DECLARATION

Name of Candidate:Alharbi, Norah Muteb S

Matric No: WOA160017

Name of Degree: Master of Computer Science (Applied Computing)

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"): Detecting Arabic Terrorism Messages on Twitter Using Machine Learning

Field of Study: Human Computer Interaction

Field of Study:

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date: 14/09/2019

Subscribed and solemnly declared before,

Witness's Signature Date:

Name:

Designation:

UNIVERSITI MALAYA

PERAKUAN KEASLIAN PENULISAN

Nama: Alharbi, Norah Muteb S

No. Matrik: WOA160017

Nama Ijazah: Sarjana Sains Komputer (Pengkomputeran Gunaan)

Tajuk Kertas Projek/Laporan Penyelidikan/Disertasi/Tesis ("Hasil Kerja ini"): Mengenalkan Pelanggaran Arab Pelanggan Pada Twitter Menggunakan Pembelajaran Machine

Bidang Penyelidikan: Interaksi Komputer Manusia

Saya dengan sesungguhnya dan sebenarnya mengaku bahawa:

(1) Saya adalah satu-satunya pengarang/penulis Hasil Kerja ini;
(2) Hasil Kerja ini adalah asli;
(3) Apa-apa penggunaan mana-mana hasil kerja yang mengandungi hakcipta telah dilakukan secara urusan yang wajar dan bagi maksud yang dibenarkan dan apa-apa petikan, ekstrak, rujukan atau pengeluaran semula daripada atau kepada mana-mana hasil kerja yang mengandungi hakcipta telah dinyatakan dengan sejelasnya dan secukupnya dan satu pengiktirafan tajuk hasil kerja tersebut dan pengarang/penulisnya telah dilakukan di dalam Hasil Kerja ini;
(4) Saya tidak mempunyai apa-apa pengetahuan sebenar atau patut semunasabahnya tahu bahawa penghasilan Hasil Kerja ini melanggar suatu hakcipta hasil kerja yang lain;
(5) Saya dengan ini menyerahkan kesemua dan tiap-tiap hak yang terkandung di dalam hakcipta Hasil Kerja ini kepada Universiti Malaya ("UM") yang seterusnya mula dari sekarang adalah tuan punya kepada hakcipta di dalam Hasil Kerja ini dan apa-apa pengeluaran semula atau penggunaan dalam apa jua bentuk atau dengan apa juga cara sekalipun adalah dilarang tanpa terlebih dahulu mendapat kebenaran bertulis dari UM;
(6) Saya sedar sepenuhnya sekiranya dalam masa penghasilan Hasil Kerja ini saya telah melanggar suatu hakcipta hasil kerja yang lain sama ada dengan niat atau sebaliknya, saya boleh dikenakan tindakan undang-undang atau apa-apa tindakan lain sebagaimana yang diputuskan oleh UM.

Tandatangan Calon   Tarikh:14/09/2019

Diperbuat dan sesungguhnya diakui di hadapan,

Tandatangan Saksi   Tarikh:

Nama:

Jawatan:

DETECTING ARABIC TERRORISM MESSAGES ON TWITTER USING MACHINE LEARNING

ABSTRACT

Terrorist groups like ISIS are spreading online propaganda using numerous social media platforms such as Twitter and Facebook. Radical groups in the Arab world are making use of these platforms at alarming rates. One of the more common approaches for stopping the use of social media by these terrorist groups involves suspending their accounts once they are discovered. However, the use of this approach requires that analysts manually read and analyze social media activities, which often involve the manual analysis of significant amounts of information. In addition, the existing works are not efficient enough to stop these malicious activities due to lack of research on the retrieval and data mining of data in Arabic, especially those involved in terrorist activities. This research is undertaken to propose an effective text classifier based on machine learning model for detecting terrorism in Twitter, it is an attempt at using a machine learning model that will automatically detect Arabic tweets from terrorist groups on the Twitter platform. Machine learning was used to aid in both the detection and categorization of a set of diverse Arabic tweets. These tweets were finally classified as either radical or not radical. This work has investigated the use of use of five text classifiers , due to the variety of philosophies behind each method and its learning process, to select the best classifier for our features. These classifiers are the Support Vector Machine (SVM), AdaBoost (discrete), AdaBoost (real), Logistic Regression and Naïve Bayes. The performance of these models was evaluated using precision, recall, F-measure, and accuracy. The work produced promising results that suggest that the use of machine learning models to detect radical Arabic content on social media platforms may be used with great potential for yielding results. The experimental results and data suggest that the use of these models yield high accuracy, with the linear classifier yielding the best results with 99.7% accuracy.

Keywords: Machine learning, Twitter, Radical, Arabic, Social Media.

MENGENAL PELANGGARAN ARAB PELANGGAN PADA TWITTER MENGGUNAKAN PEMBELAJARAN MESIN

ABSTRAK

Kumpulan pengganas seperti ISIS kini menyebarkan propaganda maya menerusi pelbagai platform media sosial seperti Twitter dan Facebook. Penggunaan platform aeperti ini oleh kumpulan radikal di negara-negara Arab adalah pada kadar yang sangat membimbangkan. Salah satu kaedah biasa untuk menghentikan penggunaan media social oleh kumpulan pengganas ini termasuklah menutup akaun media social mereka jika dikesan. Walaubagaimanapun, penggunaan kaedah ini memerlukan penganalisa meneliti dan menganalisa aktiviti media sosial yang melibatkan jumlah maklumat yang terlalu banyak. Selain itu juga, langkah sedia ada tidak cukup berkesan untuk memberhentikan aktiviti dengan niat tidak baik ini berikutan kurangnya kajian mengenai pengumpulan dan penyimpanan data melibatkan Bahasa Arab, terutamanya yang melibatkan kegiatan pengganas. Langkah ini merupakan percubaan untuk menggunakan sistem komputer (machine learning) yang secara automatiknya membantu mengenalpasti hantaran dalam Bahasa Arab yang menggunakan platform Twitter oleh kumpulan pengganas, mengkategorikan pelbagai hantaran Twitter ini kemudiannya mengklasifikasikannya sebagai radikal mahupun tidak. Proses kerja ini menggunapakai 5 pengelas perkataan berikutan kepelbagaian falsafah disebalik setiap kaedah dan proses pembelajaran masing-masing ; pengelas ini adalah Support Vector Machine (SVM), AdaBoost (discrete, real), Logistic Regression dan Naïve Bayes. Prestasi bagi setiap model ini di ukur menggunakan precision (tahap ketepatan antara maklumat yang diperlukan oleh pengguna dengan hasil daripada sistem), recall (tahap kebolehan sistem untuk mengingat kembali sesuatu maklumat, F-measure (formula ketepatan dan diukur melalui elemen precision dan recall) dan accuracy (tahap persamaan di antara nilai yang di anggarkan dan nilai sebenar. Hasil menunjukkan penggunaan sistem machine learning untuk mengenalpasti kandungan Bahasa arab yang radikal di platfom media social berpontensi untuk digunakan bagi membuahkan hasil yang lebih baik. Hasil kajian menunjukkan penggunaan model ini menunjukkan ketepatan yang lebih tinggi, terutamanya berkaitan pengelas linear membabitkan accuracy dengan peratusan sebanyak 99.7%.

Kata kunci: Machine learning, Twitter, Radikal, Bahasa Arab, Media Sosial

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

API	:	application programming interfaces
FN	:	number of false negatives
FP	:	number of false positives
IDF	:	Inverse Document Frequency
ISIS	:	Islamic State of Iraq and Syria
LR	:	Logistic Regression
ML	:	Machine Learning
NB	:	Naïve Bayes
S	:	Is referred to sample while the numbers are represented the n-gram models that are used.
S1	:	Is refers to sample of the uni-gram in the language models.
S2	:	Is refers to sample of the bi-gram in the language models.
S3	:	Is refers to sample of the ti-gram in the language models.
SA	:	Sentiment Analysis
SVM	:	Support Vector Machine
TN	:	number of number of true negatives
TP	:	number of true positives
TW-CON	:	It is the dataset containing tweets that are against ISIS and terrorism.
TW-PRO	:	Is the dataset containing tweets that support ISIS and terrorism

LIST OF APPENDICES

CHAPTER 1: INTRODUCTION

1.1 Chapter Overview

This chapter provides an introduction to the subject of the research and the research problem. In addition, this chapter presents the objectives, and discusses the scope and the motivations for this study.

1.2 Introduction

Until 2016, there were more than 3 billion internet users and this number is believed to have doubled by now (Yong, Gates, & Harrison, 2016). The introduction and application of internet has greatly revolutionized certain activities such as interactions, sales and promotion, which have recently been disrupted by the activities of electronic commerce, crypto currency, business intelligence and other information systems. Education sectors have also been affected by such reforms with the introduction of online learning platforms such as MOOC. Other sectors such as advertising, publishing, human resources, customer services and numerous other sectors are also experiencing their share of the revolution. The creation, access, the processing and the use of information have not been spared by this revolution as the internet has greatly facilitated several evolving and emerging channels of information, networking and interaction which defy the status quo (Yong, Gates, & Harrison, 2016).

People, businesses and organizations are now benefitting and enjoying the effects of the internet revolution, although this does not spare it from some of its negative impact. However, decision makers have been found to adopt and use these media and information from various online and terrestrial sources to help them make effective and efficient decisions and develop policies that will enable them to react to situations and events appropriately (Yates & Paquette, 2010).

Studies have also proven that internet technology is beleaguered by various demerits which include online scamming, spamming and fraud (Moon et al., 2010). Primarily, certain groups of criminals have also taken advantage of technology to perpetrate and execute several evil deeds. Communication and networking platforms, especially the social media platforms such as WhatsApp, Facebook, Instagram and Twitter (Kwon, Chadha, & Pellizzaro, 2017) have in no limited measure benefitted individuals and groups, whether as a means to promote their products and services, disseminate information, concepts or ideas. Twitter is one of the most used social media applications after Facebook and the fastest growing micro blogging and online social networking site that individuals and organizations adopt for publicizing and promoting their brands and disseminate information. Twitter has been found to have contributed to landmark events such as the presidential election, global monitoring, health issues, sports, health pandemics and natural and man-made disasters as such events are happening in real time (Signorini, 2011). Sometimes certain information seems to find their way to twitter before the mainstream media.

These communication platforms, especially twitter, are equally believed to have been adopted and serve as a perfect avenue for numerous terrorist establishments to spread information about their ideas which have greatly assisted their executions (Witmer and Bertram, 2016). Terrorists have taken advantage of these platforms that no study has been found to analyze their hate speeches, panic, and terror tweets. As a result, they have greatly executed numerous inhuman activities through Twitter coordinated trail (Beninati, 2016). Jihadist groups, and specifically ISIS (Islamic State of Iraq and Syria), have been able to maintain a persistent online presence by sharing content through a broad network of ”media mujahedeen” in one of the clearest incarnations of net war since it was first envisaged (Fisher, Prucha, Kaati, Omer, & Prucha, 2015).

Twitter users are referred to as tweep while the messages posted on twitter platforms are known as tweets. Tweets have been adopted by twitter users to report everything ranging from daily occurrence to the latest and global news and events (Ateefeh & Khreich, 2015). It has also been established that citizens adopt twitter as a platform to report and comment on issues of varied importance. Twitter, as a social media platform as mentioned earlier, has been embraced by many terrorist organizations to disseminate information just like other individuals and organizations and also to spread hate speeches and propaganda. The tweets of propaganda by terrorist organizations sometimes come in the form of virtual messages, videos, audios, and presentations, among others. All these are intended to help brainwash and allure their audience. Internet has been found to be an avenue used by some of the senior extremists as "battlefield for jihad, a place for missionary work, a field for confronting the enemies of God" (Ashcroft, Fisher, Kaati, Omer & Prucha, 2015).

Tweets by members of these groups are used as orders and instructions to execute various inhumane activities such as killing innocent souls, massive destruction, causing disruption and security threats, and disorder to individual lives and properties. Twitter plays an important role in enhancing the terrorists' level and increasing the number of victims, as happened in the terrorist attack in Mumbai where Twitter played an important role in enhancing terrorists' level and increasing the number of victims (Agrawal & Rao, 2011). Atefeh and Khreich's (2015) study has monitored, analyzed and processed user-generated content on social networking sites which has yielded unprecedented and valuable information which enabled individuals or organizations to implement and acquire actionable knowledge and insights. Also, their study has surveyed techniques for event detections from twitter streams. These techniques were aimed at finding real-world occurrences that unfold over space and time.

Arabic is counted among the world's most popular languages and also one among the six official languages of the United Nations organization where a considerable work has been done to develop the multilingual United Nations Bibliographic (Brahmi, Ech-Cherif, & Benyettou, 2012). Although Arabic is a widely used language, there is still little research done on retrieval or data mining (Al-Qatab & Ainon, 2010). With the presence of the effect of twitter on attacks which have claimed lives and property in the Arab world, this work intends to use Machine Learning approach to classify terror messages in the Arabic language on Twitter.

Among the world's most popular languages, Arabic is counted as one, and Arabic ranks as one of the six United Nations Organization's official languages. This was done considerably so as to develop a multilingual Bibliography of the United Nations (Ech-Cherif, Brahmi &, Benyettou. (2012). A considerable amount of research has not been conducted on data mining or retrieval considering that Arabic is one of the languages that is widely used (Al-Qatab & Aion, 2010).

This report is outlined as follows. Chapter 2 describes terrorism and provides a brief overview about machine learning techniques, and also presents related work that has been done in the area. Chapter 3 shows the stages of classification and how the classifiers were built. Chapter 4 presents the results of the experiments of this study. The discussion will be in Chapter 5 and Chapter 6 presents suggestions for future research.

1.3 Problem Statement

Studies by Witmer and Bertram (2016) and Beninati (2016) have clearly demonstrated that terrorist accessibility and connections to online materials and resources have played a significant role in their radicalization process and procedure. These terrorist groups primarily focused on destabilizing people by causing unease, fear, and terror among the majority of harmless citizens through their dissemination of ferocious audios and videos displaying fights and the massacring process of individuals who have different ideologies from the terrorist groups (Omer, 2015).

The activities and practices of terrorists have been discovered and there is evidence that they are rampant in most Arab countries (Alsaedi, 2017). The majority of these terror groups such as al-Qaeda, ISIS, and Al-Shabaab have been traced to the Arab world.

Moreover, Twitter has been discovered to be one of the most vital tools adopted by these terrorist groups to disseminate threats, enlist members, and serve as avenues for the training of new members (Gates & Podder, 2015).

One of the key techniques to help the society is by tracking and suspending accounts used by the radical groups to disseminate information about recruitment and activities, propaganda, hate speeches and other terrorist and extremist tendencies (Oh, Agrawal, & Rao, 2011). However, the method requires activities of human analysts to carefully and physically read and scrutinize the vast amount of data on the social media. Also with the huge amount of data on the Twitter platforms that is gross, rigid, abstracted, ununderstandable and unreadable, this led to an attempt to automatically detect tweets with terrorist content.

There are several English sentiment analysis datasets including terrorism detection. However, the datasets in the Arabic language are limited. The Arabic language has extended attributes with different characteristics when compared with Western languages. In addition, there are several extended Arabic social features such as culture, race and education. The Arabic social characteristics make it difficult to be used across language techniques to detect a topic like terrorism in tweets using machine learning techniques (Alshari et al, 2017). There is some limited literature on the problem of radical discovery on Arab social media. Magdy et al. (2015).

In this study, Arabic radical content on Twitter is detected and Arabic tweet is classified as supporting radicalism on twitter or no by Machine learning . Also, it focuses on classifying the text, using linguistic heuristics (language model) in the Arabic language. The ways on how to trend and implement the techniques to detect the tweets will be stated more clearly in the methodology section of the study.

Motivations of the Study

The activities and terrorist-related materials which are accessible online makes it is easy for extremist groups like ISIS to continue to explore and use Twitter as a mode for facilitating and implementing their terror activities: “Such processes frequently goes along with the transformation of trainees into persons determined to act with violence centered on extremist principles” (Omer, 2015, p. 1). As a result, many people have ended up losing their lives and properties, especially in the Arab countries, where terrorist activities are in a bigger number compared to other countries.

Unfortunately, event detection in Arabic is yet to receive sufficient attention. Although Arabic is a widely used language, there is still little research done on retrieval or data mining (Al-Qatab & Ainon, 2010). Thus, this study attempted to detect Arabic tweets that endorse radical opinions through their tweets or interaction on the platforms. This may also serve as a contribution to law enforcement agencies to be able to track the terrorists using these tweets.

Despite the huge amount of data on the Twitter platforms, these data are gross, rigid, abstracted, ununderstandable and unreadable. Thus, it is difficult to determine the columns, type, location, fields and method of extracting significant data.

1.4 Aim of Work

The principal aim of this study was to propose an effective text classifier based on machine learning model for detecting terrorism on Twitter through focusing on the twitter feature selection. Although the contribution lies in applying machine learning in a new area, the Arabic language has complex morphology and structure and it is rich in idioms and there are no articles applied to it in terrorism according to the knowledge from the LR. Therefore the n-gram models are used as feature extraction for the Arabic language in the twitter and thus it was compared with TFIDF as a benchmark. This will help analysts in their work to detect Arabic tweets that endorse radical opinions and serve as surveillance for law enforcement agencies to be able to track terrorists using these tweets.

1.5 Objectives

The overall objectives of this study were:

1. To implement a set comprehensive features that can provide an accurate terrorism detection model for Twitter platforms.
2. To develop a machine learning model using the proposed features for detecting terrorism detection on Twitter.
3. To evaluate the performance of the proposed machine learning model using a wide range of evaluation metrics (precision, recall, F-measure, and accuracy).

1.6 Research Questions

The following research questions were formulated for this study:

1. How to implement a set comprehensive features that can provide an accurate terrorism detection model for Twitter platforms?

2. How to build a machine learning model using the proposed features for detecting terrorism detection on Twitter?
3. How to evaluate the performance of the proposed machine learn model using a wide range of evaluation metrics (precision, recall, F-measure, and accuracy)?

1.7 Contributions of the Research

- **To implement a set comprehensive features and run a set of experiments to select the best machine learning**: The aim of this work is to use a machine learning model that will automatically detect Arabic tweets from terrorist groups on the Twitter platform. This study has investigated the use of five text classifiers, due to the variety of philosophies behind each method and its learning process, and selected the best classifier for the features.
- **Effective model is built for** terrorism detection: the final outcomes of this research is a machine learning model that can detect radical Arabic content on social media platforms and has a great potential for yielding results.

1.8 Scope of Research

- This study provided Arabic tweets terrorism detection model.
- It focused on the Arabic tweets as Arabic has different features such as race, education, position and culture.
- In addition, the approach employed statistical machine learning for extracting more features and relations between tweets to mine the terrorism tweets and tweeters.
- Scrawling and API techniques were used to collect the Arabic tweets from the Twitter social media.
- Statistical approaches were used to evaluate and label the dataset.

CHAPTER 2: LITERATURE REVIEW

2.1 Chapter Overview

This was the most important phase of the research that had to be conducted in order to identify the research gaps. During this phase of the research, several relevant studies were reviewed. In addition, in this chapter, a clarification of the framework of the main concepts of the study is provided.

2.2 Sentiment Analysis

Sentiment analysis is also referred to as opinion mining. It engages the application of biometrics to systematically track natural language dispensation, text analysis, and computational dialectology so as to ascertain, extract, enumerate, and study sentimental states as well as subjective data (Ravi & Ravi, 2015). Sentiment analysis is broadly employed to a voice of the client materials like reviews, in addition to survey reactions, social media and online and healthcare tools for use that vary from publicizing to customer service and medical medicine (Khan, Asghar, Ahmad, & Kundi, 2014). Sentiment analysis defines the boldness of a writer, narrator, or any other subject regarding certain topics or even the general circumstantial polarity or even emotional interaction or event. In matters pertaining to terrorism, it is believed that sentiments on Twitter would reveal rich information on the worldviews and beliefs of Tweeters.

There are a number of studies that have investigated sentiments on Twitter (Agarwal, Xie, Vovsha, Rambow, & Passonneau, 2011; Iskandar, 2017). These studies typically have the aim of categorizing the Tweets as either negative or positive or even neutral. A pipeline of classifiers employing an amalgam arrangement can enhance the precision of microblog organization, according to Khan et al. (2014). After using preprocessing approaches, the input is passed via diverse classifiers such as SWN-built classifiers. In the end, a greater accuracy is achieved compared to the techniques given for comparison. Nonetheless, additional enhancement could be generated by integrating domain-precise informal words (slang) and a lengthier set of emoticons. Prieto-Merino et al. (2014) recommended a wordlist-centric technique that chains diverse vocabularies as well as lexicons in order to investigate the sentiments of messages. The major stress was on the precise cataloguing of sentiments concerning dialect in the Tweets. Their projected approach has diverse components called subjectivity detection, Tweet seizing and riddling, as well as sentiment notching. It is evident that these components are backed by diverse lexicons such as the view lexicon and emoticon fonts. They attained 92% accurateness in twofold classification and over 85% in multi-course classification (Prieto-Merino et al., 2014).

The work of Khan et al. (2014) regarding sentiment arrangements recommended a lexicon-centered approach to excerpts, preprocess, and pigeonhole operator sentiments from online societies like the terror groups. They employed diverse dictionaries counting SWN as well as user-distinct lexicons to define the polarity scores regarding sentiments of words. Furthermore, domain-explicit terms were pulled out to enhance reliability. The main confines of their technique encompassed lack of management multifarious sentences and a dialect and ironic decrees, which, if integrated, could lead to enhanced sentiment classification. Prieto-Merino et al. (2014) composed Tweets in line with diverse enquiry terms such as despair, eating, and flu. They used machine learning algorithms for arrangement and feature assortment to check public apprehension in Spain and Portugal. As a result, they obtained F-measure values of about 0.8 and 0.9 that are significant, likened to the model approaches. Nonetheless, the system relies on the labeled training dataset and there was no upkeep for sorting of dialect, dictionary, and sphere-reliant terms in various fields (Prieto-Merino et al., 2014).

The unverified learning method is an additional approach to sentiment scrutiny of data on Twitter (Montejo-Ráez et al., 2012). As a way of computing sentiment scores at the global level, Montejo-Ráez et al. (2012) employed SWN, as well as an unsystematic walk approach to analyze the weighting of Tweets. It is apparent that the projected algorithm was not reliable about the labeled training dataset and it has showed key enhancement compared to the baseline techniques. Nonetheless, the limitations encompassed a lack of negation management, discrepancies, and labor-intensive notes in Tweets in the totalling of the closing soppiness notch. The internet has a significant effect on the precision of Twitter-built sentiment scrutiny usages. Kundi et al. (2014) offered a structure to sense and score the dialects in Tweets by applying diverse wordlists such as SWN amongst other sentiment possessions. The researchers realized improved outcomes likened to the baseline techniques. The principal confines of their work comprised an inadequate focus on managing emoticons and the necessity for further class setting-cognizant and sentiment-profound enchantment rectification elements.

Furthermore, Kundi et al. (2014) failed to look at domain-precision confrontations, which brand the arrangement less efficiently. Ranco et al. (2015) argue that Twitter has a substantial effect on typical market rates. To examine the Tweets, the "occasion study" monetary method was embraced for involuntary documentation of happenings as Twitter-centered capacity heaps. The "occasion study" monetary method helps in scrutinizing the positive and negative sentiments articulated during the heaps. Lastly, “events study” was employed to recognize the connection between Tweets and stock morals (Ranco et al., 2015). The chief drawback of the exertion was the shortage of emoticons and dialect elements for further precise sentiment sorting of Tweets in the stock market forecast. In Ribeiro et al. (2015), an integrated method for executing Tweet-grounded sentiment analysis was presented. The suggested technique encompassed four components: data collection, noise reduction, lexicon generation and sentiment classification.

They presented four central algorithms to execute the above-mentioned modules. The outcomes acquired from the tests which were run on the iPhone 6 dataset established that the projected method is more efficient than the common techniques (Ribeiro et al., 2015). Nevertheless, to realize more values, additional trials were necessitated on bigger datasets with Tweet grinding in spilling approach (Kundi et al., 2014).

Tang et al. (2015) recommended a broadcast-centric sentiment analysis method for Twitter. The approach was anticipated to incorporate diverse emotional inklings into a fused ideal and trains on both labelled and unlabeled datasets by converting the propagation sensation interchangeably. The trials conducted on multiple datasets validated the efficiency of the projected approach. The suggested technique is centered on general-drive learning and could be heightened to a pigeonhole domain-precise words in diverse spheres. Tang et al.'s (2015) work regarding contextual analysis recommended a lexicon-improved polarity classification approach to work out contextual polarity at diverse levels.

Khan et al. (2014) provided a lexicon reinforced tactic for Twitter-grounded data scrutiny that seizes the sentiment category of words in different backgrounds and appraise the sentiment notches as required. Their method was built on co-rate word signs in diverse spheres at both Tweet and unit heights. The projected technique was assessed employing three datasets and realized 4-5% greater precision than the evaluation techniques. Nevertheless, the structure was not supplemented with emoticons as well as dialect arrangements. According to Khan et al. (2014), a convolutional multiple kernel learning-centered approach pertaining to sentiment analysis was presented. The analysis encompassed various forms of multimedia content such as audio, text, and video clips. Topographies were removed from the multimedia content by using commencement ideals in the innermost coat of a profound convolutional neural system ideal. As such, the outcomes revealed that an enhancement of around 14% was realized more than the model approaches.

To sum up, it is evident that the prevailing techniques for sentiment cataloguing are reliant on the accessibility of big glossed datasets, which leads to performance deprivation. The finding is according to recent studies. The difficulty with unsupervised approaches is that they rely on various aspects:

a. Reliance upon widely existing sentiment lexicons; and
b. Their application of rudimentary sentiment lexicons with restricted backing for slang, emoticons, and field-precise words.

For this reason, a further arrangement technique is necessitated that could further successfully pigeonhole Tweets (Prieto-Merino et al., 2014). As a result, work is required towards designing an effective Twitter-centric sentiment analysis arrangement employing an amalgam scheme for classification with improved set of emoticons, slang words as well as pipeline of sentiment classifiers with the aptitude of precisely sort general-drive and domain-explicit opinion words while concurrently realizing more-vigorous fallouts that are equivalent to the performance outcomes of supervised and unsubstantiated tactics.

In a recent study by Agarwal et al. (2011), the use of sentiment analysis of Tweets was conducted. The authors categorized each Tweet as neutral, positive, or negative. In addition, certain features were centered on the schism of words. The centering was resolved by employing numerous lexicons, for instance, the Dictionary of Affect in

Language (DAL). From their experiment employing an SVM as well as the classifier with unigram model, they obtained 71.36% precision.

2.2.1 Social Media

Social media encompasses a collection of internet-grounded applications that allow users to create and share content and/or take part in social networking. Twitter, which allows the users to send and read about 280-characters which forms a Tweet, is among the most recognized micro-blogs. Twitter has been selected for the current research due to its informal nature, which provides bloggers the ability and freedom to employ casual language and free expressions in their tweets. Another reason for its selection is its message conciseness, with the limited length of 280 characters per post renders (Duwairi et al , 2014).

Extremist groups have used Twitter in the furtherance of terrorist acts. As described by Agrawal and Rao (2011) in numerous cases, persons and organizations employ social media to draw attention to the battalions and fundraisers to precise causes (the original causes which drove them to Twitter). Terrorists (militants), who are individuals, who contribute to terrorism, have greatly expanded their application of Twitter, besides other social media platforms like Facebook and YouTube (Omer, 2015).

Moreover, the use of hashtags also provides a fertile area for investigation. Hashtags are used to mark diverse themes and topics in Tweets, for instance, #Terrorism attack in Syria (Omer, 2015). The symbols are usually applied to enhance the visibility of the tweet.

2.2.2 Radical Groups and the Application of Social Media

Social media is employed to communicate with acquaintances and household members, but it is also a method for promoting personal opinions. Sometimes radical, even dangerous, views are disseminated through the platform. While community organizations use social media to promote local sports and fundraising events, business interests and other social activities, militants have been found to employ Twitter to recruit members for subversive movements. Omer (2015) believes that persons associated with Jihad movements have increased their use of Twitter and other platforms as a means to enhance organization. In 2015, about 90,000 Twitter accounts were reported to have supported radical organizations according to the scholar.

ISIS (the Islamic State of Iraq and Syria), for instance, employs dispersed arrangements of network structures and plans to spread vivid video content from the battleground in close to real time (Omer, 2015).

ISIS and similar groups are able to defy conventional media channels and promote activities considered criminal by international standards. ISIS has effectively applied social media to hire new militants globally, according to reports (Kwon et al., 2017). A rebel leader, sitting in front of a computer or using a smartphone, can promote or even coordinate subversive activities in another part of the world. For instance, mujahedeen have used the media to broadcast executions of hostages and prisoners (Kwon et al., 2017). One of their aids was their Twitter guide providing information in real-time.

As mentioned, one of the globally recognized terrorist groups is ISIS. The present work contends that posts transmitted online by user groups with well-known ISIS connections are likely to be disseminating posts with terror content (Gates & Podder, 2015). Underlying this thesis is an assumption that there are Tweets that are posted and support ISIS' worldview as well as those that are in fact anti-ISIS (Omer, 2015). Some of the pro-ISIS Tweet's hashtags include #ISLAMICSTATE. Chatfield et al. (2015) also found that a strong link is recognized in the Arabic world between terrorist activities, particularly recruitment, and the use of Twitter accounts.

2.2.3 Terrorism

Terror is a term that grew to prominence during the times of the French Revolution during a period of mass executions conducted by the revolutionaries (Rapport, 2015). In more recent times, terrorism has become understood as the intentional application of indiscriminate violence for the purposes of creating fear among people to achieve political aims (Fortna, 2015). Terrorism can be top-down, when caused by states, or bottom-up, when carried out by non-state groups (Rapport, 2015). Numerous criminal groups have profited from the global spread, advancement, and speed of the internet. Exploiting the internet has placed radical groups in a position to spread their ideology, and broaden their reach leading to a situation where they have the opportunity to hire individuals globally. Additionally, certain social media platforms have provided rogues with a media network to publicize their posts and radicalize people. An earlier study by Torok (2013) revealed that the use of the internet by radical organizations has experienced phenomenal growth recently (Torok , 2013).

According to Blaker (2015), since 9/11, terrorist groups such as Al-Qaeda have exploited the internet to efficiently and reasonably communicate, publicize propaganda, and hire enthusiasts with little jeopardy of retaliation from the counterterrorism organizations. Furthermore, Weimann (2015) argued that during the early establishment of the internet as the environment of choice, the application of password-secured conversation boards permitted prevailing affiliates to communicate amongst themselves with relative anonymity. Additionally, discussion boards permitted members of Islamic terror cells to talk on all topics linked with Jihad. The year 2008 is seen as a turning point in which radical terrorist groups became many times more prevalent online (Weidmann, 2015).

The same author argued that dependence on Twitter to disseminate radical thoughts has been particularly high due to the strong results that such groups have achieved (Weidmann, 2015). The scholar noted that recruitment has been positively influenced by content on social media. The levels of sophistication have also increased in terms of radical group planning and social media. Blaker (2015) argued that groups such as Jabhat al-Nusra, Islamic State, and an associate of al Qaeda based in Syria, have had a pronounced success in their social media campaigns and have moved on to generate devoted media wings. Gates and Podder (2015) stated that in the case of ISIS, media content was particularly impressive.

There are videos that have been released by its media arm, Al Hayat, that show different sides of the militant crew. Nevertheless, it shows the face of tension and terror as it displays such things as decapitated heads held by children. Worse still, there are shared videos that are Western-friendly as they showcase such things as Nutella jars being held by IS militants as they pose with them to show how familiar they are with the culture and lifestyle of the West. Thus, the terror organizations have been showing social media content not only of violence but also of humanistic events and happenings. Moreover, Gates and Podder (2015) note that the social media content of ISIS is focused on re-building the area.

There are so many products of propaganda but most of them focus on the providing new construction, justice and governance. The issue of legitimacy is very important. These departments of social media experts focus on the construction of propaganda that can be leveraged on Twitter and their equivalents and the involvement of their cohorts. Stern and Berger (2015) noted some of the video editing and even cinematography that had been applied to terrorist cell recruitment programs. The pair maintained that via the application of battleground footage, music videos, and bilingual forms, groups are in a position to spread their ideology to a broader audience addressee than what was conceivable in the earlier decade. Indeed, even government bodies in Western settings have recognized a tougher situation in restricting participation by their citizens in organizations such as ISIS. Blaker (2015) asserted that the IS (The Islamic State's) has specifically validated the ability for enlightening a large community of supporters via their application of Twitter that has surpassed the aptitude of other jihadist groups. Blaker (2015) argued that IS's propaganda segment, Al-Hayat, generates high-quality staffing videos that impersonate the distinct impacts shown in action movies, accompanied by music videos, as well as an online magazine. Moreover, Stern and Berger (2015) affirmed that the content of their propaganda frequently encompasses a collocation of dangerous ferocity with acts of a utopic society.

Stern and Berger (2015) affirmed that pictures and videos portraying decapitations, implementations, and battlefield footage substitute contrary to phantasmagorias of IS-run nursing households and the construction of city substructure. The IS propaganda helps to further their ideology while still appealing to a range of audience affiliates. Athough it is apparent that discussions take place on Twitter from mundane life happenings to exhaustive dialogues of matters on their relevant group's ideology (Blaker, 2015), it is common for extremist groups to affirm battlefield conquests or to assert a terrorist outbreak on the social media. The IS habitually escorts their posts with pictures of the demise, torment, and the implementations of those who are regarded to be nonbelievers (Stern & Berger, 2015). It is apparent that the exhibitionism of ferocity established in the continuous stream of propaganda on the social media is among one of the features that delineate the IS' propaganda plan from other extremist groups (Stern & Berger, 2015). As outlined by the researchers, the violence demonstrated towards Westerners as well as Muslims alike occurs all through the bombast of the IS and was partly accredited to the institution's split-up from Al Qaeda in 2014 (Stern & Berger, 2015).

The IS' partiality for vehement battlefield theatrics, in addition to implementations, could be likened to Al Qaeda's past propaganda that is focused on spiritual teachings as well as beliefs (Stern & Berger, 2015). Consequently, this plan led to a rise in teenagers who showed enthusiasm towards the support of the IS, compared to an adventure, instead of an earnest religious ideology (Stern & Berger, 2015).

Stern and Berger (2015) affirmed that the Zora Foundation was intended to be effortlessly reachable mainly on Facebook and Twitter focusing on organizing probable female recruits of the ISIS for their Jihad. It is apparent that although Al Qaeda depended on conventional human interaction in crucial locations like religious centers, learning institutions, and marketplaces to assist in their recruitment efforts, they did not re-appropriate this approach in their social media campaign (Weimann, 2015). Nevertheless, the IS had emboldened those who are radicalized, with some living beyond the IS-controlled terrain, to dynamically involve with women with whom they are networked on the social media (Hosken, 2015).

Iskandar (2017) conducted a comparative study sentiment analysis technique to improvise the current sentiment analysis techniques to discover the undertakings of terrorism further precisely. The scholar compared the sentences that have been categorized into positive, negative and neutral categories to the earlier sentences of a certain account possessor centered on the soppiness notch for the newest and prior sentences. Moreover, Iskandar (2017) the sentiment scores are calculated by using top 50 radicalism keywords itemized in the US Department of Homeland Security (DHS) for social media (2011). In contrast, the relations between keywords in DHS and the words that did not have synonymy in Tweet are omitted.

Kwon et al. (2017) described how terrorism has increased in the social media. They used geographic, social and temporal dimensions as features to predict the tweets of terrorism. These tweets are chosen randomly. Bodine-Baron et al. (2016) explained how the ISIS has been employed by the social media platforms to spread the ideas within positive covers in twitter. Furthermore, the researchers tried to comprehend how the ISIS gets supporters from the social media, through employing lexical analysis approaches. The contemporary increase and territorial advances of the ISIS (Islamic State) have indicated substantial interest in the organization. Abdulla (2007) proves that such militant organizations are typically political instead of being religious.

2.3 Sentiment Analysis Datasets

There is a wide corpus of literature leveraging social networks as a foundation of data with diverse methods (Cuesta, Barrero, & Moreno, 2014). The network configuration lends itself to display analysis as well as expert studies, which have been piloted on an all-encompassing multifaceted network description that permits profound scrutiny of Twitter dealings as well their designs. A social community's conduct could be evaluated from freely accessible data, for instance, geographical and topological properties as well as their associated conduct. Additionally, other graph-centric methods are to model, in addition to simulating the network (Cuesta et al., 2014). It is apparent that traditional Twitter application results in sharp eruptions of activity when activities occur (Cuesta et al., 2014). The activity is referred to formally in Twitter's network as "Trending Topics" or "TTs". Developing topics could be identified in real time examining these eruptions and likening them with the prior undertaking. Another study by Cuesta et al. (2014) demonstrates how to inevitably determine remarkable topics from these developing topics, taking into consideration the poster's reliability grounded on a number of topographies, for example, user behavior, message content and topic-centered extents.

2.3.1 Twitter Dataset

Such approaches result in interesting outcomes such as real-time event discovery. For example, Cuesta et al. (2014) were successful in detecting earthquakes bearing in mind Twitter users as "sensors" realizing great outcomes. Furthermore, Twitter could be employed as a public health sign, which permits tracking diseases over time and space, evaluating risk aspects and medication applications. As a domain of study, it joins the methods of ordinary linguistic dispensation, computational languages and text mining. The form of analysis gives a great supply for end-users. For instance, the method is relevant for product appraisal summarization, which could assist consumers for shopping online. According to Chae (2015), corporations and organizations could similarly influence their usefulness as enormous information basis on which to approximate the view created by their merchandise or services. Moreover, it could be employed by politicians as well as policymakers to scrutinize public opinions.

According to Chae (2015), there are certain interests from investigators as a result of its open nature and the likelihood of gathering mass opinion through sentiment-forte discovery. Agarwal and Sureka (2015) projected a technique for involuntary collection of messages with positive characteristics besides negative characteristic traits like sentiments, on which they execute language scrutiny and generate a sentiment classifier. The most current method uses semantic analysis to extract and categorize the sentiment linked to distinct entities. The method can help to deduce not only the sentiment as a dichotomist unit, but also extract a scaled outcome. The shortcoming of this is that the broader the scale, the less accuracy the outcomes.

2.3.2 Twitter Data Extraction and Analysis Framework

The aim is to offer an easy-to-apply Twitter data extraction and scrutiny for research drives. Consequently, the platform is segmental, with numerous autonomous modules implemented as diverse programs. Adding novel models or modifying them is straightforward, making it simpler to adopt the platform to the necessities of the investigator (Agarwal & Sureka, 2015). The fundamental piece in this design is a database that maintains the Tweets mined by the use and makes them accessible for further dispensation. There are various layers of processing and these modules require swapping data amid them, applying open data formats such as JSON. By default, the framework encompasses modules for creation report as well as sentiment analysis (Agarwal & Sureka, 2015). The platform is separated into various components:

2.3.2.1 Miner

This is the core of the platform. It listens indeterminately to a filtered Twitter stream as well as store the entire status apprises into the database. The mining module backs three forms of operation: single mode, parallel mode, and In-serial mode (Chae, 2015).

2.3.2.2 Classifier

It is evident that a Web interface permits turncoats to assist with the administered classification of Tweets (Cuesta et al., 2014). As a result, the collaborators must agree on the Tweet's sentiment and categorize it as required.

2.3.2.3 Trainer

It encompasses simplified interfaces NLTK library that aids in forming corpora and models further than Tweet collections.

2.3.2.4 Tester

It encompasses a collection of tools that assist evaluating the skilled model's aptness, counting implements for manual classification of Tweets through CLI in addition to web interfaces, cataloging of a Twitter creek in real-time as well as cross-authentication of produced classifiers (Chae, 2015).

2.3.2.5 Reports generator

The report generation component helps in amassing numbers from the database mutually via sentiment and quantitative variables. Statistical or numbers' analysis is executed, applying a collection of scripts inscribed in R. As a way of decoupling the statistical analysis from the backend, a midway module produces a collection of CSV records from the database. As a result, this module is in a position to work on quantitative reports, offering certain rudimentary statistics, in addition to sentiment reports, with the aim of classifying the sentiments (Agarwal & Sureka, 2015). It is apparent that MongoDB database would be a suitable choice for fast write, particularly for fast document writing because of its atomic representation.

Data extraction, as well as sentiment analysis, is shared into three phases: training, data acquisition, training for sentiment analysis as well as report creation (Cuesta et al., 2014). The initial phase concerns congregating data from Twitter with the Miner. After that, the processing of training the classifier is done and the sentiment analysis is undertaken. Finally, the platform forms a collection of reports, counting the sentiment analysis if it is empowered.

a. Datasets

The process of scrutinizing structured data has been employed over the last decade. The traditional Relational Database Management System (RDBMS) has been a leading method to handle the data. With the rising volumes of unstructured data on numerous sources such as social media, the worldwide web, and blog data that are regarded as Big Data, a distinct computer processor is limited in its ability to process such enormous volume of data (Chae, 2015). For this reason, the RDBMS fails to handle the unstructured data; a non-ancient database is necessitated to manoeuvre the data, which is regarded as the NoSQL database.

b. Data Retrieval

Before repossessing the data, certain queries ought to be sorted out concerning data characteristics. For example, if it is static, why it is crucial, how to employ it, and what volume is needed. It is essential to comprehend that tracking a particular keyword on a hashtag instead of the one not attached is vital (Cuesta et al., 2014). Twitter-API is a broadly employed application to recover, read, and write Twitter data.

c. Ranking and Classifying Twitter Users

There are diverse categories of user's networks; for example, a network of users in a certain event or hashtag, an account, and a group.

2.4 Feature Extraction and Machine Learning

The current operation of separation of pro-ISIS Tweets and Tweets contrary to ISIS would be engaged. Also, a list of operators who are separated into groups with recognizable ISIS enthusiasts would be involved (Bjoergum, 2014). A research using sentiment analysis was steered; as a result, they categorized a Tweet as being positive, neutral, or negative. Certain features were centered on the divergence of words. This is usually resolved by applying various phrasebooks, for example, WordNet. In their experiment employing an SVM classifier and unigram topographies, they attained 71.36% exactness (Magdy, Darwish, & Abokhodair, 2015). Then again, when unigram topographies are combined with semi-features, the outcome rises to about 75.39% indicating the input of the semi-features for Tweets sentiment classification.

According to Bird et al. (2008), even though the majority of the undeveloped states are not excused from civil turbulence and war, the map shows that radicalism is widely common in developed democracies besides the Middle East. A shared outlook is that the culprits of radicalism are often spiritual radicals aiming at those with differing principles. Ellis (2003) argues that by 1995, more than 50% of terrorist groups were centered on religion. Additionally, Crenshaw (2001) claims that terrorism ought to be perceived as a tactical reaction to the American power in the perspective of a worldwide civil war. Fanatical religious philosophies play a vital role in encouraging terrorism. Blomberg (2008) mentions that authors such as Wintrobe (2002) and Bernholz (2004) have likewise studied the obligation of augmented fundamentalism as well as group cohesion in making terrorist undertakings.

2.4.1 Feature Extraction

The authors applied specific feature extraction modules that are incorporated to account for the linguistic features of Arabic. The result sensitivity was useful, but the key disadvantage was the thrilling lack of preprocessing, which is certainly vital for Arabic texts in Almas et al. (2007). Omer (2015) applied several machine learning such as AdaBoost, SVM, and Naive Bayes to extract the terrorism messages in Twitter by statistical gain algorithm, stylometric analysis, writer invariant method, and kernel trick technique. They applied these machine learning classifiers on just English messages Tweets without the labeling of radical Twitter accounts.

2.4.1.1 Language Model

A statistical language model or a language model is the distribution of words using probability over sequences. It is essential in natural language processing application as it provides a way to estimate the frequency of different phrases, especially those that give text as an output. The application of language modelling can be seen in machine translation, speech recognition, parsing, handwriting recognition, speech recognition, information retrieval, part of speech tagging among many other uses. However, the problem of data sparsity is a significant hindrance to the creation of a functional language model as most of the probable word sequence cannot be observed during training. The solution which has been presented for this issue is to assume that the possibility of a word solely depends on the number (n) of words that appeared previously. This solution is called the unigram model or the n-gram where n =1.

The n-gram model or unigram model is a language model used in predicting the likely item or number in a sequence using the formula a (n-1) – order. Markov model and the n-gram model are now widely accepted as the best language model for communication theory, probability and computational linguistics (for example in the processing of natural statistical language). Simplicity and scalability with larger n are the clear benefits of the n-gram model and all other algorithms that use it. A great model with a well-designed space and time tradeoff can store massive contents thereby making it easy for small experiments to scale up without any problems. The n-gram is a sequence of a specified number (n) items from a particular speech or text such as syllabus, base pairing, phonemes, words and letters depending on the application. With the use of Latin numeral prefixes, an n-gram with a size of 1 is called unigram, a size of 2 is called diagram or bigram, and trigram for three and so on.

According to Posadas-Duran et al. (2015), a syntactic n-gram is of different kinds based on the type of information used in their design (POS tags, lemmas, relations or Words). There is a significant difference between traditional n-gram and the syntactic n-gram, and this can be seen in their formation. Syntactic n-gram follows syntactic relations from the syntactic trees while the traditional n-gram is built from words surface string just as they appear in the text. Also, this difference is what makes the syntactic n-gram produce better results than the traditional n-gram especially through the impact it has on text classification. The syntactic n-gram model depends on the syntactic-based n-gram features to determine the age, personality traits and gender of the creator of a given text. In the experiment, only three results were achieved for three languages which showed a high performance. However, to make the approach better, a weight scheme was the proposal that will help create the right balance in training data.

Ashcroft et al. (2015) adopted a machine-learning method to categorize Tweets as pro- or anti-ISIS. They focused on English Tweets that encompass a reference to a set of predefined English hash-tags linked with ISIS. The limitation of their approach is that it is highly dependent on data. The AdaBoost classifier outperformed in accuracy over other machine learning classifiers.

Yang et al. (2011) combined in their study between machine learning and semantic-oriented.The researchers confirmed that adding further feature sets (syntactic, stylistic, content-specific, and lexicon features) could meaningfully improve the efficacy of the classifiers for terrorists' views identification with the restrictions: greater quantity of physically labelled training data and time-consuming. Kwon (2017) also investigated Tweets and links to terrorism and described how terrorism increased in the social media. They used (geographic, social and temporal) dimensions as features to predict the Tweets on terrorism. The Tweets were chosen randomly.

From previous studies, it is apparent that text classification is a vital study area of text mining. The unique drive of text classification is to identify, comprehend and consolidate diverse forms of texts or documents. The overall classification tactics are perceived as administered learning, which deduces similarity amidst an assortment of classified texts for training drives. The prevailing categorization methods are perceptibly not content-oriented and inhibited at single word level (Guo, Shao, & Hua, 2010). Studies have been conducted to propose Arabic script webpage linguistic documentation scheme employing a resolution tree-ARTMAP tactic (DTA) to determine wide-ranging characteristics in a web document which will later lead to optimization performance of the language (Selamat & Ng, 2011). Another study by Brahmi et al. (2012) revealed stemming tactics for Arabic topic modelling to enable users to find out interesting topics to publish in the press. A contemporary study on language classifications have been undertaken by Guo et al. (2010) that have similarly recommended an automatic text classification applying human cognitive approach.

Moreover, the Dormant Dirichlet Allocation model was employed for the extraction of the latent topics from the three Arabic real-world corpora. In other studies, the document was vectorized, not as per the occurrence of words, but on the heart of the manifestation of security linked terms (Cheong & Lee, 2011). The occurrence of one or more applicable words to predefined classes (extremism, war-radicalism, Jihad amid others) was applied to realize the last category. It was concluded that more than 90% accuracy resulted in the idea that keywords would be a better technique for categorizing Tweets. On the individual and momentous level, they were in a position to foresee forthcoming backing or antagonism of an operator for ISIS with almost 87% exactness. Beneath this, they trained an SVM classifier with a linear kernel of evasion constraints (Klausen, 2015). One of the key hitches experienced in their operation was to detach pro-ISIS Tweets to con-ISIS ones. As a result, they realized that in anti-ISIS Tweets when denoting Islamic State operators inscribe ISIS was 77.3% of the Tweets against that of pro-ISIS Tweets that were 93.1% (Omer, 2015). The decent outcome of the classifier showed that SVM would be a good method of categorizing Tweets.

Numerous social media networks, for instance, Facebook and Twitter are operating towards maintaining these platforms clean by appending those who are encouraging vehement content or extremist conduct. Nevertheless, as a result of the volume besides the speed of the produced data, it is still puzzling to identify those unruly users precisely and in a suitable manner. The latest research has focused on exploring the online conduct of pro-extremist operators primarily by undertaking content-grounded study to detect distinctive textual features that could help in automatic discovery of these operators (Ashcroft, Fisher, Kaati, Omer, & Prucha, 2015).

Ashcroft et al. (2015) made an effort to routinely identify Jihadist messages on Twitter. As such, they employed a machine-learning technique to categorize Tweets as ISIS enthusiasts or not. They focused on English Tweets that encompass a reference to a collection of predefined English hashtags linked with ISIS. It is apparent that one of the confines of their method is that it is exceedingly reliant on the data. Choudhary et al. (2015) explored the prevailing literature on counter radicalism in addition to social network scrutiny. Some of the explored glitches in this domain are associated with ascertaining key-players, discovering conduct patterns, community unearthing, and unsettling terrorist links. Consequently, they established that the application of Social Network Analysis (SNA) is one of the greatest fruitful approaches for counter terrorism in social networking.

According to Choi et al. (2012), the data set of n-gram is also used in the classification documents as it is a co-occurrence statistics data collector. As a co-occurrence statistics data collector, it collects adjacent texts from the training manuals due to its linguistic benefits. These linguistic benefits allow it to extract core features from the training manuals written in the natural language. N-grams are used in filling incomplete sentences, speech recognition and topic discovery. N-grams are usually extracted in two ways; the first is the extraction carried out on every adjacent English character from the document, while the second is the extraction carried out on every adjacent word. However, there are a few issues associated with the first extraction method, and that is, the size of the n-gram is usually very large and takes a lot of time to process. The second method works best with diagrams and trigrams and to effectively detect threats from terrorist's texts and tweets, trigram, bigram and unigram data frequency was built and developed from a set of words used in the classification of such tweets.

2.4.1.2 Algorithms

Support Vector Machine (SVM) is a large margin classifier. It has the objective to discover a frontier between two classes that exploit the space between data shared as demonstrated in the diagram below (Omer, 2015). The Support Vector Machine (SVM) are new promising non-parametric, non-linier techniques of classifications, which have in the past indicated good results in the optimal character recognition, medical diagnosis as well as forecasting among other fields (Auria & Moro, 2008). The importance of SVM classification technique is application in solvency analysis where it is used in developing a function, which can separate accurately the insolvent and solvent companies' space, through score value benchmarking. SVM have been considered by Auria and Moro (2008) as a new technique that is mainly suitable for tasks of binary classification, relating to, and containing, elements of non-parametric applied statistics, machine learning and neural networks.

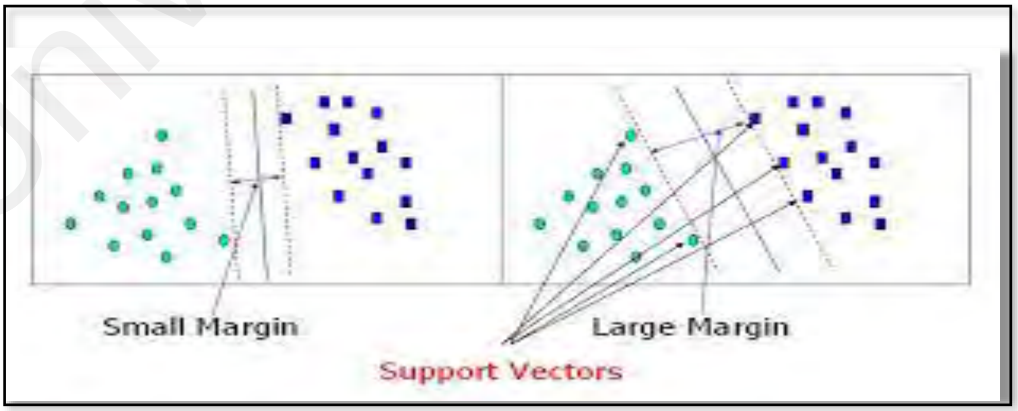

Figure 2.1: Support Vector Machine (Omer, 2015)

Figure 2.1 outlines margin and support vectors. The aim of the SVM is usually to discover the hyperplane that provides the largest space to the training instances (van Ginkel, 2015). Text classifying deals with very many features. Since SVMs use overfitting protection, which does not necessarily depend on the number of features, they have the potential to handle these large feature spaces. One of the greatest attributes of the SVMs is their ability to learn fast. This gives it the advantage of being independent of the space features and dimensionality. The complexity of the hypothesis is weighed by SVMs depending on the data separating margin and not the feature's number.

AdaBoost is one of the machine learning algorithms that is grounded on boosting. Importantly, boosting is a technique that syndicates discreet imprecise guidelines to generate a perfect classifier. It is centered on the supposition that it is easier to discover numerous rules rather than, one precise rule (Collins et al., 2002; Chen et al., 2014). Adaboost (discrete), on the other hand, is found by Viola and Jones, (2004) to be an important classification technique because it handles real-valued hypothesis under the convex maximization auspices. The advantage and importance of this boosting algorithm in supervised learning are many such as training the moderately accurate learning, acquiring its hypothesis that is weak and combining them so that a strong classifier can be outputted which ensures that the accuracy is boosted up to high levels arbitrary. Discrete Adboots also has been seen by Nishii and Eguchi (2005) to be among the most popular boosting algorithm provable, and its importance is its ability to use weak hypothesis with restricted output to discrete classes sets that are combined by its vial linear vote leveraging coefficient. The theoretical support has been strongly advocated for discrete Adaboost due to its advantage of handling real-valued weak hypothesis. Unlike discrete Adaboost, the real Adaboost is found by Nock and Nielson (2007) to handle arbitrary weak hypothesis in real value. This makes the weak hypothesis leveraging hypothesis to differ in output. The importance and benefit of real Adaboost is improving radically the weak leaving performance by building smaller number of classifiers that are weak iteratively, and further making sure that they are fussed to strong ones. Specifically, the real Adaboost can adapt to a problem by combining together the weak classifiers to make them strong. Real Adaboost deals with mapping of weak classifiers confidence to a real valued prediction from sample space instead of Boolean prediction. There is an advantage of real-valued weaker classifier compared to the Boolean ones in its ability to discriminate (Jin et al., 2009). Although the discrete Adaboost trains the final classifier, similar accuracy is achieved when using Adaboost mostly. However, the former has much more classifiers compared to the latter. Real Adaboost namely has the ability of achieving accuracy that is high as a number of weak classifiers are used to fix it, exactly in the cases of learning online.

Further, one of the several methods of machine learning is Logistic Regression which is a type of regression analysis used for predicting dichotomous dependent variables by estimating the probability of the event's occurrence (Omer, 2015). It works by taking inputs and multiplying the value of input with value of weight (Indra et al., 2016). This is one of the key classifiers that learns from the input what features are the most important and useful and thereafter discriminate them between different possible categories. The discrimination model of Logistic Regression implies the computing of P (y|x) through discrimination among various values based on the given input x in class y. The equation can be written as shown below.

$$P(c|x) = \sum_{i}^{N} = 1\, W_i f_i$$

The $P(c|x)$ value is not calculated actually directly by the use of the previous formula due to the fact that it will result in value from -∞ to ∞ which implies that it cannot provide an output between the value of 0 and 1. In generating output that ranges between 0 and 1, the following function is utilized (Indra et al., 2016).

$$P(c|x) = \frac{1}{z} \exp \sum_{i} W_i f_i$$

Naive Bayes is a probabilistic classifier based on applying Bayes theorem assuming independence between the features (Omer, 2015). Naive Bayes computes the probability p of a document d being of class c: p (c|d). Given a document d to be classified, represented by a vector d = (d1... Dn) (Rish, 2001, August). The Bayer classifiers usually assign the most likely class to a distributed example described by its feature vector. Naive Bayes is remarkably productive in apply in is application typically competitive with far more refined techniques (Rish, 2001, August). Naive Bayes has been proven to be useful in many practical applications which include text classification, management of system performance and in medical diagnosis. There are two instances where the naive Bayer works best, and they are: as completely independent features (just as expected) and functionally dependent features (not as apparent as the other instance) while reaching its worst performance between these extremes (Mujtaba et al. 2017). Naive Bayer is one of the popular inductive learning classifiers in supervised machine learning classifiers and is considered a reliable and efficient decision model. NB, which is derived from the Bayes' theorem with strong independence assumptions among its other features, is also straightforward to use, quick, and frequently produces better accuracy compared with other classifiers (Mujtaba et al. 2017).

2.4.2 Machine Learning

With the rapid growth of online information, text categorization has become one of the key techniques for handling and organizing text data. Machine learning is regarded as the knowledge that discovers how algorithms could be created for the purposes of learning from data and generates forthcoming forecasts (Fuchs, 2017). Algorithms be constructing a mathematical ideal centered on certain example involvements and apply the ideal to create particular estimates or pronouncements. These algorithms build a mathematical model on the basis of some example inputs and utilizes the model for predictions or decision-making. It is through the processes of employing the model that any input could be drawn to an array (Omer, 2015).

A machine learning algorithm is critical in training models that involves individuals' engagement or aid. Rather than static programming that directs the computer on the operation to undertake, machine learning algorithm would build systems that could learn from information amassed together (van Ginkel, 2015). Through machine learning, learning techniques are typically categorized into three groups: reinforcement, unsupervised, and supervised learning. The goal of machine learning is to learn a model from training data that allows us to make predictions about future data. For machine learning to work, the algorithm needs to be fed with input data. The cleaned data is split into two datasets, training and testing dataset. The training data is used for training and developing the model. Finally, the model is then used for the test set to evaluate performance (Raschka, 2015).

The outcomes of experiments are envisioned applying a confusion matrix or contingency table. It is employed in machine learning to envisage the outcomes of an administered learning algorithm (Omer, 2015). The process entails two columns as well as rows which surround the sum of factual positives, true negatives, false negatives, and false positive cases. The following shows the structure of a contingency table.

Table 2.1: The structure of a contingency

	P' (Predicted)	n' (Predicted)
P (Actual)	Positive(True)	Negative(False)
n (Actual)	Positive(False)	Negative(True)

As a result of the succession of terrorist assaults as well as the development of IS's power internationally, it has become essential to develop novel ways for further operative elucidations that could inhibit attacks, categorize defendants, respond to the moment, and similarly accelerate studies (Ferrara, Wang, Varol, Flammini, & Galstyan, 2016). Consequently, technology could assist in guarding against terrorism. Unfortunately, one thing that makes IS so challenging to fight is because of the terrorist set-up (network), which is dispersed with small cells of operators spread globally. This makes it problematic for law enforcement entities to forecast where a certain extremist organization would attack next. For this reason, anti-terrorist consultants progressively prefer Artificial Intelligence (AI), as well as Machine Learning methods, to mine large amounts of data and discover inconsistencies in doubtful conditions. Currently, AI is being used in an effort to halt extremists. As such, researchers are evolving a varied range of systems as a way of enhancing security.

Developing computational tools encompasses modernized cameras to track suspects, image corresponding algorithms to match suspects on user-uploaded videos and photos with openly acknowledged extremist faces, besides bright sensors to read the iris of each voyager (terrorist) and evade deceitful characteristics (Ferrara et al., 2016). Additionally, frequently when terrorists are preparing to carry out a radical assault they often depend on social networks, encrypted usages, and private emissary applications to harmonize their attacks. As a result, Facebook is constructing "text-centered" from previously eradicated posts that prized or braced extremist organizations (Ferrara et al., 2016). The distinctive action feeds those signals into a machine-learning structure, which after some time, would learn how to identify similar posts. Likewise, Google is aiming at detestable content with machine learning-built approaches operating together with human critics as well as Non-governmental Organizations (NGOs) in an effort to host a nuanced tactic to expurgate terrorist media.

Nonetheless, researchers are confronting a life-threatening absurdity in that extremists themselves could apply such identical technologies. Thus, it may have become a fundamental race for an invention which necessitates a more practical method. Instead of acclimatizing technologies to remain ahead of developing hazards and fluctuating strategies, there is a demand to be ahead of the extremists and grow "overmatching" security arrangements that safeguard the public, their autonomy and give consent to travel and business.

Ferrara et al. (2016) present a machine learning structure that controls a collection of metadata, network, and chronological topographies to discover terrorist users, and forecast adopters as well as contact mutuality on social media. Ferrara et al. (2016) have exploited an exceptional dataset containing numerous Tweets engendered by over 25 thousand users who have been physically recognized, testified, and suspended by Twitter as a result of their engagement with terrorist operations. Furthermore, the platforms control millions of tweets created by a haphazard taster of 25 thousand consistent users who were open to, or inspired, by radical content. Consequently, two predicting tasks have been executed:

i. To identify radical users.

ii. To approximate whether unvarying users would assume the terrorist content.

To foresee if users would react to exchanges introduced by radicals, it was important that all anticipating tasks are set up in duo scenarios; they include a post hoc forecast task and a simulated real-time forecast task. The performance of this framework is encouraging, yielding to the diverse predicting setups up to 93% AUC for terrorist user discovery which equals to 80% of AUC for content espousal forecast, and lasts up to 72% of AUC for contact reciprocity projecting. For this reason, it could be concluded that providing an exhaustive feature analysis which aids and defines the developing signals that give projecting power in diverse cases is of great significance. Two pertinent study trends have been developed in the computational social sciences as well as in the computer science groups. According to Ferrara et al. (2016), computational research focuses widely on comprehending the social marvel circling around radical propaganda employing online data as a deputation to investigate group as well as individual conduct. Recently, numerous studies have focused on English besides Arabic audiences online to research the impact of ISIS' propaganda in addition to its radicalization. The majority of the ISIS' accomplishments on Twitter is because of a limited number of highly-operative accounts (Ferrara et al. 2016). The analysis demonstrates that definitely, an imperfect number of ISIS accounts realized a very great discernibility and followership.

Johansson, Kaati, and Sahlgren (2016) affirm that the aptitude to spread information promptly through huge geographical areas makes the internet a vital expediter in the radicalization course and arrangements for extremist assaults. The situation could mutually be a challenge and an asset for security institutions. One of the key challenges for security institutions is the absolute extent of data obtainable on the Internet. It is unbearable for human predictors to read via all that is written online. The idea to notice terrorism-allied content on the Internet is not renewed, as it is supposed that the detection of terrorist activities on the Internet would help prevent the influence of looming radicalism (Johansson et al., 2016). Essentially, there are present-day calls for study ventures by the European Research (EU) as well as innovation program, Horizon 20201, whose purpose is to identify and scrutinize terrorism-linked content on the Internet with the aim of combatting radicalism. According to Johansson et al. (2016), data mining, as well as machine learning algorithms, are crucial in classifying the written content of websites as either terror-linked or vice versa.

A technique that permits the managing or handling of internet service providers' traffic does it in the correct time, permitting for extensive monitoring. Nonetheless, this could be juxtaposed with the method presented by Johansson et al. (2016) whereby a much smaller subgroup of websites, which are connected to recognized radical websites, is subject to scrutiny. Furthermore, it is in the latter methodology recommended that only user-created content on the sites is investigated, not the web traffic per se. It is apparent that this is an imperative transformation as the methodology employed in the prototype arrangement in analyzing what individuals write about. In contrast, another research by Johansson et al. (2016) asserts that terror-linked content is normally perceived by terrorists as well as their supporters and could be employed by data mining implemented to study a "Typical-Terror-Conduct". Nevertheless, the outlook is challenging as the majority of the persons who are not terrorists would be concerned with reading the content regarding similar topics as radicals.

Safekeeping agents and intelligence specialists have reported concern for radicalism on the web and can be classified as ‘suspected terrorists' employing a comprehensive methodology. O'Hara and Stevens (2015, p. 402) argue that radicalism is “a belief/behavior nexus in considerable tension with the embedding society” and that “‘extremism’ [is] the violent pursuit of radical goals” and that both were being enabled through the internet. It is argued by radicals that a vital role is played by networks in issuing mechanisms to actualize their objectives. The existence of a central cluster or a clique makes it more important as there is interconnection between a wider group and a subgroup, and as a result most people in the subgroup will be known by almost everyone in the subgroup. Communications from sources that are heterodox should easily be cut and instead, better connections to orthodox adherents be provided (O'Hara & Stevens, 2015, p. 412).

Investigators claim that the search for radicals within a group is an expensive endeavor. Moreover, the risks are doubled when considering the costs of missing a single terrorist in the haystack of non-terrorist users (Johansson et al., 2016).

2.5 Research Gap:

Sentiment analysis is broadly employed to a voice of the client materials like reviews, in addition to survey reactions, social media and online and healthcare tools for use that vary from publicizing to customer service and medical medicine (Khan, Asghar, Ahmad, & Kundi, 2014). In matters pertaining to terrorism, it is believed that sentiments on Twitter would reveal rich information on the worldviews and beliefs of Tweeters. There are a number of studies that have investigated sentiments on Twitter (Agarwal et al., 2011; Iskandar, 2017). These studies typically have the aim of categorizing the Tweets as either negative or positive or even neutral. Prieto-Merino et al. (2014) recommended a wordlist-centric technique that chains diverse vocabularies as well as lexicons in order to investigate the sentiments of messages. The work of Khan et al. (2014) regarding sentiment arrangements recommended a lexicon-centered approach to excerpts, preprocess, and pigeonhole operator sentiments from online societies like the terror groups. Prieto-Merino et al. (2014) composed Tweets in line with diverse enquiry terms such as despair, eating, and flu. They used machine learning algorithms for arrangement and feature assortment to check public apprehension in Spain and Portugal. The unverified learning method is an additional approach to sentiment scrutiny of data on Twitter (Montejo et al., 2012).

There is some limited literature on the problem of radical detection on Arabic social media.Magdy et al. (2015) studied the classifier of SVM that detects whether or not a user is a supporter or an opposition of ISIS. This is made possible by the text features displayed in a tweet by the user. Most of the tweets are distinguished at this stage and the declaration of their support. Some instances on twitter are examples that are used to detect jihadist supporters by studying their lingual and temporal features. Mubarak et al. (2017) proposed an approach that that could easily expand and create a highlight of obscene phrases that are later used to detect tweets that are profane by list .

Radical speech has been investigated quite extensively in English social media content. Cohen et al. (2014) attempts to find lone wolf terrorism by detecting various traits of it such as leakage, fixation, and identification warning behaviors through text analysis. The authors mentioned various NLP based methods such as Part of speech, lemmatization of words, frequent combinations of certain key terms, usage of positive adjectives as their primary means of text analysis. Chatfield et al. (2015) conducted trend analysis, social network analysis and content analysis of a total of 3,039 tweets posted by one of the verified information disseminator for the IS cause. Using social network analysis, they showed that there are four distinct clusters of Twitter users: (1) international media, (2) regional Arabic media, (3) IS sympathizers and (4) IS fighters. Through content analysis, they classify the tweet contents into four categories: propaganda, radicalization, terrorist recruitment and others.

Scanlon et al. (2014) presented supervised learning and natural language processing methods for automatically identifying forum posts intended to recruit new violent extremist members from social media websites. They used data from the western jihadist website Ansar AL Jihad Network, which was compiled by the University of Arizona’s Dark Web Project. Multiple judges manually annotated a sample of these data, marking 192 randomly sampled posts as recruiting (YES) or non-recruiting (NO). The authors experimented with naive Bayes models, logistic regression, classification trees, boosting, and support vector machines (SVM) to classify the forum posts. Evaluation with ROC curves shows that SVM classifier achieves an 89% area under the curve (AUC), a significant improvement over the 63% AUC performance achieved by simplest naive Bayes model.

Kalpakis et al. (2018) compared various characteristics of suspended Twitter accounts against those of non-suspended accounts to determine the key factors that are capable of providing weak signals for distinguishing ordinary Twitter users from those with subversive behavior based on the analysis of a variety of textual, spatial, temporal and social network features. The authors took into account various factors such as the lifetime of suspended accounts, user accounts from the social network perspective (i.e. based on their connectivity with other user accounts) and geolocation information extracted from the textual content of user posts.

Sentiment Analysis of Twitter Data, Agarwal et al. (2011) introduced a novel POS-specific prior polarity features for classifying tweets into positive, negative and neutral classes. They created a total of 100 hand crafted features. The features are broadly categorized into positive integer type, real number or Boolean type. Each of these broad categories is further broken down into polar and non-polar subcategories. Finally they used SVM with 5-fold cross validation for classification purpose. In their experiments, they show that the hybrid model of unigram with these 100 features provide the best classification performance.

Iskandar (2017) conducted a comparative study sentiment analysis technique to improvise the current sentiment analysis techniques to discover the undertakings of terrorism further precisely. The scholar compared the sentences that have been categorized into positive, negative and neutral categories to the earlier sentences of a certain account possessor centered on the soppiness notch for the newest and prior sentences. Moreover, Iskandar (2017) the sentiment scores are calculated by using top

50 radicalism keywords itemized in the US Department of Homeland Security (DHS) for social media (2011).

Kwon (2017) described how terrorism has increased in the social media. They used geographic, social and temporal dimensions as features to predict the tweets of terrorism. These tweets are chosen randomly. A research using sentiment analysis was steered; as a result, they categorized a Tweet as being positive, neutral, or negative. Certain features were centered on the divergence of words (Bjoergum, 2014). In their experiment employing an SVM classifier and unigram topographies, they attained 71.36% exactness (Magdy, Darwish, & Abokhodair, 2015). Then again, when unigram topographies are combined with semi-features, the outcome rises to about 75.39% indicating the input of the semi-features for Tweets sentiment classification.

The authors applied specific feature extraction modules that are incorporated to account for the linguistic features of Arabic. The result sensitivity was useful, but the key disadvantage was the thrilling lack of preprocessing, which is certainly vital for Arabic texts in (Almas et al., 2007). Omer (2015) applied several machine learning such as AdaBoost, SVM, and Naive Bayes to extract the terrorism messages in Twitter by statistical gain algorithm, stylometric analysis, writer invariant method, and kernel trick technique. They applied these machine learning classifiers on just English messages Tweets without the labeling of radical Twitter accounts.

Ashcroft et al. (2015) adopted a machine-learning method to categorize Tweets as pro- or anti-ISIS. They focused on English Tweets that encompass a reference to a set of predefined English hash-tags linked with ISIS. The limitation of their approach is that it is highly dependent on data. The AdaBoost classifier outperformed in accuracy over other machine learning classifiers. Yang et al. (2011) combined in their study between machine learning and semantic-oriented.The researchers confirmed that adding further feature sets (syntactic, stylistic, content-specific, and lexicon features) could meaningfully improve the efficacy of the classifiers for terrorists' views identification with the restrictions: greater quantity of physically labelled training data and time-consuming.

The latest research has focused on exploring the online conduct of pro-extremist operators primarily by undertaking content-grounded study to detect distinctive textual features that could help in automatic discovery of these operators (Ashcroft, Fisher, Kaati, Omer, & Prucha, 2015). Choudhary et al. (2015) explored the prevailing literature on counter radicalism in addition to social network scrutiny. Some of the explored glitches in this domain are associated with ascertaining key-players, discovering conduct patterns, community unearthing, and unsettling terrorist links. Consequently, they established that the application of Social Network Analysis (SNA) is one of the greatest fruitful approaches for counter terrorism in social networking.

A machine learning algorithm is critical in training models that involves individuals' engagement or aid. Rather than static programming that directs the computer on the operation to undertake, machine learning algorithm would build systems that could learn from information amassed together (van Ginkel, 2015). The latest research has focused on exploring the online conduct of pro-extremist operators primarily by undertaking content-grounded study to detect distinctive textual features that could help in automatic discovery of these operators (Ashcroft, Fisher, Kaati, Omer, & Prucha, 2015). Choudhary et al. (2015) explored the prevailing literature on counter radicalism in addition to social network scrutiny. Some of the explored glitches in this domain are associated with ascertaining key-players, discovering conduct patterns, community unearthing, and unsettling terrorist links. Consequently, they established that the

application of Social Network Analysis (SNA) is one of the greatest fruitful approaches for counter terrorism in social networking.

A machine learning algorithm is critical in training models that involves individuals' engagement or aid. Rather than static programming that directs the computer on the operation to undertake, machine learning algorithm would build systems that could learn from information amassed together (van Ginkel, 2015).

There has been an investigation of radical speech in quite an extensive way in regard to English content of social media. Character n-grams were observed as improved features of prediction than n-gram words by Waseem and Hovy (2016) for detecting sexist and racist tweets. They also argued that using gender as a feature of addition could only yield little improvement to results of grouping while a decrease in performance would be registered by additional location information. Large margins that used Boosted gradient decision trees outperformed models based on n-gram classifications. Embeddings that used Long Short-Term memory network were studied by classifiers. An approach was proposed by a recent study to equipp with lexical and syntactic features to detect offensive language. It was also to display precision of a higher level compared to learning-based approach that is traditional (Chen et al., 2012).

SVM was applied for the purpose of detecting cyber bullying and determined that incorporating user-based content improved the detection accuracy of SVM (Dadvar et al., 2013). In a recent study by Agarwal et al. (2011), the use of sentiment analysis of Tweets was conducted. The authors categorized each Tweet as neutral, positive, or negative. In addition, certain features were centered on the schism of words. The centering was resolved by employing numerous lexicons, for instance, the Dictionary of Affect in Language (DAL). From their experiment employing an SVM as well as the classifier with unigram model, they obtained 71.36% precision.

2.6 Summaries of the related work:

There is some limited literature on the problem of radical detection on Arabic social media. Radical speech has been investigated quite extensively in English social media content. While there is a lot of literature in the field of sentiment analysis (SA). The following table rviewed some studies related to this research.

Table 2.2 : A summary of related work

References	Problem Statement	Method/Techniques	Contribution	Limitations/ weak
Magdy et al., 2015	Better understanding of the roots of ISIS organization and its supporters.	SVM Bag-of-words, F1 score	-Determine distinguishing language patterns to indicate current support or opposition for ISIS -Predict future support or opposition of ISIS - Identify potential reasons for subsequent support of ISIS and the sources of opposition to the group.	-It can only find ideological supporters, not the actual fighters. -Twitter actively suspends extremist accounts. Also twitter has a lot of restrictions such as only last 3200 tweets of a user can be retrieved. This can be termed as limitation since a lot of valuable data can not be accessed and used in the algorithm.
Chin et al., 2016	To assess the public sentiment for the 2016 presidential candidates	SVM Naive Byes.	-Various ways to analyze the sentiments of the general public. -Twitter was chosen as a channel to get the views of people.	The results are only limited to Twitter users
Varol et al., 2017	The social media users have to experience so many information campaigns on a daily, which play their part in shaping and changing public opinion	-K-Nearest Neighbor with Dynamic Time Warping (KNN-DTW.	The objective : -Certain topics getting as top trends on social media are fabricated, -It is important to understand that nefarious purposes can be behind these trends like propaganda of terrorists	The work was just based onTwitter Analysis and these results cannot be generalized to general social media population
Rauchfleisch et al., 2017	How they verify the information coming from social media	-API of Twitter relevant tweets. -The logistic	-Aanalyse the source as well as verify the UGC	Various other methods that journalists can use

References	Problem Statement	Method/Techniques	Contribution	Limitations/ weak
	users	regression model (LR).	collected by journalists How the information was actually sourced and later verified to be shared with the general public can be explored	to collect and verify information.
Gilbert, N., & Huang, R. (2013).	Sarcasm is something which is used by individuals to amuse, mock or insult, but these sarcastic approach can be manipulated with different methods, so it is important to know that what was the sarcastic approach behind certain thing so that any misinterpretation can be avoided,	- API of Twitter . - SVM techniques	The solution: it was analyzed that how sarcasm is considered or understood by general public	There was no other source to validate what was actually intended by the Twitter users.
Carchiolo et al., 2015	The Diseases around the world are a common problem, and it is quite important to know how dynamics of diseases are perceived by the individuals	- SNOMED-CT terminology to get clear terms for health topics. -The logistic regression (LR) .	-Twitter has been used as a great source, so researchers also used Twitter for this health-related topic to get solution for the mentioned problem .	This study has been limited to certain geographical area as it collected tweets within the New York region.
Aramaki et al., 2011	The social media and internet comes with so much information on a daily basis, but the question is whether this information can be useful and authentic to be associated with the real world	-Support Vector Machine (SVM).	Analyze how Twitter and information coming from Twitter can be useful to detect diseases or epidemic such as Influenza	The real world information from Twitter should have been more diversified and expanded
Neethu, S., & Rajasree, R., 2013	-Used symbolic techniques based on the knowledge approach and machine learning techniques. -Sentiment analysis are used to determine the levels of topics on the twitter .	- SVM, -naïve Bayes (NB). .	There are some specific feature extraction methods that can be used to extract relevant information. -Standard database is not available to analyze the twitter sentiment.	The research was limited to140 characters in both knowledge base approach and machine learning approach.
Caplan, J. 2013	The goal of the study was to get proper insights into the process in which	-Coding techniques	The aim was to measure evaluation in the informal, political exchange	The research was limited to analysing the use of twitter in the election

References	Problem Statement	Method/Techniques	Contribution	Limitations/ weak
	candidates in the 2nd Congressional District of Virginia used twitter to reveal the post information.		and the social influence on the targeted numerous sectors of American society.	campaign of Republican Congressman and Democratic candidate.
Mittal, A., & Goel, A., 2012	The research used sentiment analysis and machine learning process. In the research correlation was determined to use the twitter data and prediction of public mood.	- Linear regression - Logistics regression - SVMs (LIBSVM).	The hypothesis was based on the premise of behavioral economics and it analysed the impact of moods and emotions on the decision making process. The research determined the co-relation between the market sentiment and the public sentiment.	The research was limited to two parameters for the optimal parametrization. In the research, large enough test sets were used but limited data was used with 30 to 40 entries.
Benhardus, J., 2013	In the research outline methods were used to detect and identify the trending topics for data streaming. The research used data from twitter streaming and then analysed trending topics on the twitter.	-The term frequency inverse document frequency analysis and the relative normalized term frequency analysis . - NPL	-Term frequency inverse document was used to identify the trending topics.	The limited stream data was used for the garden hose streaming and twitter activity was limited to 15% only. The possible extensions are limited for maximum improvement and it can be used only for the limited topics .
Sun et al., 2003	Use of information extraction in the research was significant for the web intelligence and the IE systems. The research considered extraction patterns with the restricted templates	The research used the IE method for SVM techniques to extract the data.	The classification model was used to develop a support for the vector machine. The IE experiments were used to evaluate the proposed methods for the text collection and the terrorism domain.	The research was limited to the extraction of perpetrator entities with the collection of untagged documents.
Sakaki et al., 2010	Twitter is now being used regularly for different awareness purposes, especially through different text updates. It is now used as an important application as an earthquake reporting system.	Probabilistic models have been used to identify the impact of earthquake and to generate the study.	Identifying the use of twitter in earthquake awareness.	-There is still limited access of different people to twitter services. -The difference between typhoons and earthquakes as they occur simultaneously.
Lee et al., 2013	The goal of geographical analysis and surveillance is to	Real-time flu and cancer surveillance system have been	Identifying the use of twitter in surveillance of	Limited knowledge about the system and limited

References	Problem Statement	Method/Techniques	Contribution	Limitations/ weak
	track the disease spread in different parts of the world by measuring the volume of flu cancer tweets generated in the region.	used.	informative data and the use of twitter in fighting against flu and cancer.	availability of time were the major limitations.
Anjaria, M. and Guddeti, R.M.R., 2014	Avalanche of messages on social networks, especially twitter, make it a very attractive medium of user data analysis through posts about consumer brands and services they use, or expression about political and religious views.	-NaIve Bayes, -SVM -MaxEnt -ANN .	-The impact of twitter in validating important political news that identify the most suitable approach for utilizing twitter features along with sentiment analysis	A large variation in results made it difficult to identify the actual results easily.
Inoshika et al., 2013	In order to create the dataset, the frequencies of words were used.	- SVM	-The impact of twitter in validating important political news which identify the most suitable approach for utilising twitter features for classification using SVM	-In order to apply machine learning, a proper feature set was required. -There was limited resources available to generate the most suitable data.
Neha Upadhyay, & Prof. Angad Singh 2016)	Symbolic methods or Knowledge base approach and Machine learning strategies are the two primary procedures used as a part of opinion analysis.	-Naïve Bayes	The concepts of sentiment analysis in detail were identified and the relation between sentiment analysis and machine learning and how it is influenced by the use of twitter	A very complex module architecture was used which required expert assistance and required different classifiers for data collection and interpretation of data
Rohit J., 2016	Sentiment analysis has become a major discussion in the field of research and technology, especially with the advent of social networking sites and large amount of data processed .	-Naïve Bayes, -SVM -MaxEnt	Studying the concept of sentiment analysis in detail with the impact of machine learning on sentiment analysis through twitter.	Limited knowledge about the system and limited availability of time were the major limitations
Tarun Mirani and S. Sasi., 2016	ISIS takes advantage of the social media to continuously communicate using coded words or to establish their indirect presence	- SITA -SVM -ME -RF -DT -BG	Study the impact of twitter in validating important political news identifying the most suitable approach for utilizing twitter .	Two step process was difficult to comprehend, especially to interpret data from different languages.
Ming Yang	Developing and	-SVM	-Identifying the	The approach

References	Problem Statement	Method/Techniques	Contribution	Limitations/ weak
et al., 2011	evaluating informatics tools and frameworks to measure the activities within the social media networks.	-Naïve Bayes	value of social media analytics in developing and evaluating the information -Identifying the role of social media in radial opinion mining .	requires large amount of data for supervised learning to be operated manually.
Enghin Omer, 2015	Internet technology comes with numerous benefits including sharing information and ideas as well as accessing them fast and easy.	-SVM -NB -AdaBoost	Identifying the impact of twitter in identifying various social issues and studying how jihadists use social media as an important medium	The method required extensive training to identify the data and search results, time and resource limitations were also observed.
Ikonomakis et al., 2005	There are multi-source data available online, and consequently have different formats, different preferred vocabularies and often significantly different writing styles even for documents within one genre.	-Decision trees -Naıve-Bayes	-Identify the impact of text classification in machine learning -Study the impact of data available online in machine learning.	Data was difficult to comprehend and required expert assistance.

2.7 Summary

This chapter reviewed several relevant studies. In addition, a clarification of the framework of the main concepts of the study was provided and it also highlighted the importance of this research.

CHAPTER 3: METHODOLOGY

3.1 Chapter Overview

This chapter discusses the the Sentiment Analysis (SA) classification process which extracts the features from Twitter for predicting the terror orientation of tweets. Besides, the phases of the SA process are able to represent the main methodology of extracting the orientation of text from twitter. The terror orientation process is divided into three main phases: data analysis, feature extraction and training model as shown in the figure below.

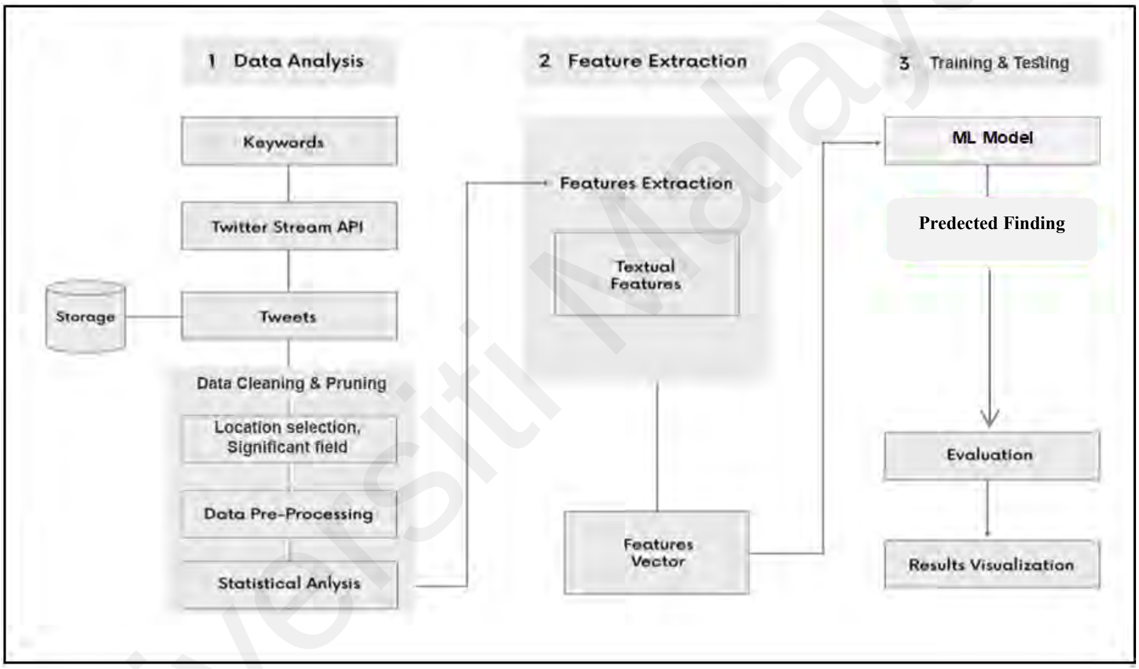

Figure 3.1: Sentiment Analysis Processes in the Proposed Approach

3.2 Data Analysis

This section shows how data was collected and divided, as well as the application of statistical methods. It also shows how Cronbach's Alpha, which is a measure of scale reliability, was calculated. The pre-processing process is also explained in this section.

3.2.1 Dataset

Twitter tracks system is adopted in the study to track words, phrases and hashtags related or often mentioned and posted in relation to "terrorism”. Data was collected using Twitter API and was further built as a data set. Twitter search has been used as it provides a search API that permits search for tweets using certain language. The language of this study will thereby be set as Arabic using syntax Arabic, which will enable access to, interaction with, and obtain Arabic tweets.

The Middle East region is defined as the spatial scope of the research under the coordinates of the length and width. The data for this study was collected during the month of March of 2018. The number of collected tweets was 135,069. However, the number of tweets was reduced to 346 tweets as a huge number of tweets was discarded and ignored in the cleaning step in the pre-processing. The data cleaning was an important process in the data collection procedure that was able to remove and prune tweets in considering several issues such as duplication, redundancy, objective text, abstraction and slang language.

It was found that there were some Tweets in languages other than Arabic. Further, there were also some Tweets that were not related to terrorism at all and they had no violent /radical content, so they were ignored. Similar and repeated tweets, which were in large numbers, were eliminated. Also, incomplete tweets were excluded. Tweets with links only were also excluded. In addition, there were tweets of news accounts that broadcast news stories about events that were not considered supportive or against terrorism. Tweets that were written in a language other than Arabic were also discarded.

In this work, the dataset collected was divided into two separate datasets. The research team termed the first dataset, TW-PRO, and the second dataset, TW-CON. TW-PRO Tweets are those from, and for, pro radical groups. In other words, the dataset has Tweets which endorse ISIS and terrorism. Examples of these tweets are:

"الهمم عالية والبسمة لا تفارق وجوهنا لا يوجد ذهاب إلى #إدلب ولا إلى #درعا نحن في #جيش_ا
نختار الذهاب إلى الجنة"

Translation: "The high motivation in fact does not leave our faces there is no going to # Idlib nor to # Daraa We are in the # Army of Islam we choose to go to paradise"

In contrast, TW-CON Tweets are those which are against ISIS and terrorism. TW-CON consists of tweets with non-radical content from different hashtags. Examples of these tweets are:

"أن تكون ضد# داعش يعني أن تكون طلقة في بندقية #التحالف_الدولي العدواني أو عصابات
#الحشد_الشعبي و#إيران فكلهم إرهاب".

Translation: "To be against the # ISIS does not mean to be shot in the gun # aggressive international coalition or gangs # Crowd_Arabic and # Iran are all terrorism".

A set of tweets containing hashtags that were related to the terrorism groups on Twitter was also collected. Examples of the hashtags that were utilized to collect the data are provided below in Table 3.1:

Table 3.1: Examples of the hashtag

Translation (English)	Arabic
# Army of Islam	#جيش_ا
#eastern Gota	#الغوطةالشرقية
#Syria	#سوريا
#Syrian_National_Army	#الجيش_الوطني_السوري
# Damascus	#دمشق
# Organization_state	#تنظيم_الدولة
#Islamic_Caliphate	#ا ة_ا ية
#Islamic_country	#الدولة_ا ية
# ISIS	#الدولة_ا ية في_العراق_والشام

TW-CON consists of tweets with non-radical content from different hashtags. Examples of tweets are:

" أن تكون ضد #داعش يعني أن تكون طائفي بتقوية #التحالف الدولي العدواني أو عصابات

#الحشد الشعبي و #إيران فكلهم إرهاب "

A total of 135608 tweets were collected and stored extracting the useful information from the json files containing the tweets. The description about the data set can be found in Table 3.2.

Table 3.2: Description of dataset

Total number of tweets	135069
Number of tweets used	346
Positive	146
Negative	200
Agree	339
Not agree	7
Language	Arabic

The information about the tweets is shown in the following table:

Table 3.3: The information about a tweet

Text	280
No. Hash	40
Location	22.91, -45.37, 2000 KM
Time	3-25/03/2018

3.2.1.1 Manual Labelling

The Arabic tweets were collected from Twitter through the number of specific hashtags that are mentioned in Appendix A and based on previous studies (Ali, 2016; Omer, 2015; Salmi Abdellatif, 2016). Then the data set was cleaned, and pre-processing applied (346 tweets). The 346 tweets were presented to four experts (native speakers of

Arabic, who are also Arabic language teachers), who classified the tweets based on the following definitions:

Definition 1: "the illegal application of force as well as violence alongside individuals or property to threaten or pressure a regime, the civilian populace, or any sector thereof, in continuance of partisan or social goals" (Ali, 2016, p. 1).

Definition 2:

القتل و الاغتيال والتخريب والتدمير ونشر الشائعات ، والتهديد وصنوف ابتزاز والاعتداء وأي نوع من اخافة يهدف الى خدمة غرض سياسي او استراتيجي" (Al-Arman & Al-Shawabkeh, 2014).

Definition 2: "Murder, assassination, vandalism, destruction, dissemination of rumors, threats, acts of extortion, assault and any kind of intimidation aimed at serving a political or strategic purpose" (Al-Arman & Al-Shawabkeh, 2014). Also, return to some of the common terminology and concepts of terrorism mentioned in the research(Salmi, 2016), such as:

(طواغيت العرب ، عملاء امريكا، الخيانة ،الولاء للبغدادي ،خليفة المسلمين،...الخ)

Translation: (Arab tyrants, American agents, betrayal, loyalty to Baghdadi, the caliph of Muslims, etc.). The label data set was used to train five classification models in the study's classification stage.

3.2.2 Preprocessing

When amassed, the data is commonly not "clean". As mentioned earlier, this is due to misspelled words, undesirable elements, and irrelevant elements that may influence the classification. For these reasons, it is essential to clean the data. The tweets format is different from the other text as it contains only 280 characters, and it includes URLs, and symbols such as “#” for hashtags, “@” for user mentions, and “RT” for re-tweet. In the first stage of data cleaning, all the URLs such as "http://example.com", #hashtags, the "@" symbol, duplicate tweets "RT" meaning retweets that wouldn't add any value to the analysis, were removed. In the second stage, the other noisy characters such as numbers, and foreign letters were removed. Once the data had been cleaned, a preprocessing was undertaken.

The preprocessing phase plays a major role in the classification task. It helps to improve the performance of the classification process by reducing the noise in the tweets. To prepare the data for building the mode, the following tasks were accomplished:

- Tokenization
- Cleaning of the the data by removing the (stop words, punctuation, blank spaces, diacritics (tashkeel), etc.

1) Tokenization

When a text is split into numbers or symbols (or words) the intervention is known as tokenization. The products are known as tokens. Tokenization is often used in lexical analysis for the purpose of exploring word in a text (Omar, 2015).

2) Removal of Stop words

This process involves removing stop words such as prepositions and pro-nouns, e.g. (هذا,في , من, أنا etc.), which do not carry any meaning or feeling and thus are not necessary. To remove stop words from tweets, a list of stop words was used. The removal of unnecessary words is an essential process in unstructured data mining because it helps to improve the classification performance and produces accurate results.

3) Removal of Punctuation and Blank Spaces

This step involves removing punctuations such as the following (. , - , / ,...) from the texts, which have no benefit and replacing them with blank spaces. Further, where there are blank spaces next to each other, these are removed to leave a single blank space.

4) Removal of diacritics (tashkeel)

A diacritic is an additional small character attached to a letter, either as a superscript or a subscript Fathah, (الحركات'" The Arabic diacritic including Al Harakaat. Kasrah, Dammah and also Sukoon), Al Tanween "التنوين" (Tanween Fath or Kasr or Damm) and Al Shaddah "الش ّ د ة". The main goal of using diacritics is to change the pronunciation of the letter which may change the word's ultimate meaning. Diacritics are frequently used in the Qur'an, other religious texts, and in Arabic language learning books. However, much of the posts that are written in Arabic online omit the tashkeel. It is only when a poem, or a religious text (such as the .Qur'an, or the Hadeeth) is quoted that the tashkeel is usually present

3.3 Feature Extraction

3.3.1 Feature Vectors

A feature vector is used to represent an object. It can be thought of as an n-dimensional vector which is based on numerical features (Omer, 2015). Feature vector creation in our work involved three textual features in language model (uni-gram, bi-gram, ti-gram). The main benefits of language model and n-gram are typically considered as scalability and simplicity. In the case of a larger n it is possible for a model to be used to store a lot of context with a space-time that is well understood. In addition, n-gram enables the scaling up of the efficiency of small experiments. In our case, unigram, bigram and trigram from the types of n-grams were used. The reasons why these were chosen, as opposed to a 4-gram or 5-gram, were time, cost and size savings (Choi et al., 2014).

Table 3.4: Size of features

Sample	No of features	Approaches
S1	3317	Benchmark
S2	8424	Proposed
S3	13628	Proposed

Table 3.4 shows three sample features that were used in this work (S1,S2,S3). "S" refers to sample while the numbers represent the n-gram models that are used. S1, S2 and S3 refer to the uni-gram, bi-gram and ti-gram in the language models respectively. For example, the tweet:

"To be against ISIS does not mean to be shot in an aggressive international coalition gun"

It is represented three times in the vocabulary of feature extraction (uni-gram, bi-gram, ti-gram). Therefore the tokenization of the vocabulary is represented as:

S1 :ان تكون ،ضد ، داع   يعني ، etc,.

Translation S1: To, be, against, ISIS, dose, not, means, etc ,

S2 : أن تكون ، أنتكون ضدتكونضد ،داعش ،ضد داع   داع   يعني ،

Translation S2: To, be against, against, against ISIS, ISIS ,…

S3 : أنتكون ،أنتكون ،ضد ، انتكونضد تكونضد ، داعش تكونضد داع   ضد داعش

، داعش يعني

Translation S3: To, be, to be, against, to be against, be against, ISIS,…

Then, the vector of features is extracted three times from the n-gram model within the frequency of each token and thus will be introduced to the classification methods.

3.3.2 Feature Extraction Vector Representation

There are several techniques to represent the features and most of these techniques is called bag-of-word, BoW or document matrix. It represents the documents as the rows of matrix and features as the columns. However, the value of a feature in the document is represented as a binary value (1 if the feature appeared in the document or 0 otherwise), and this representation has a weakness to present the frequent terms and simple semantic issues. Therefore, the weighting schema is one of the effective metrics to extract more relations between the words in the corpus or documents (Hong & Davison, 2010; Wang et al. 2012). One of the most typically applied weighing schemes is tf-idf (term frequency-inverse document frequency). The representation is applied to assess the importance of a word to a given document in a corpus of texts. Aiming to depict the Tweets' textual content, the team used a vector-space model whereby a Tweet mi is showed as a word vector (v1, v2, v3, v4 . . . , vd), and in this word vector d is the word vocabulary size and vj is the tf-idf weight of j th term in the Tweet mi. Specifically, the tf-idf weight is made up of the following terms:

1. TF (Term Frequency) is the number of incidences a word is found in a document, divided by the document's total number.
2. TF (t) = (the number of incidences term t is found in a document) / (the total terms).
3. IDF (Inverse Document Frequency): this number is determined as a logarithm of the number of the documents in a corpus divided by the number of documents where the given word appears.

IDF(t) = log_e(Document total number / the number of documents with word t).

For example:

Assume there are 1000 documents which collectively feature the term "kill" 10 times (10 instances in total), and one document contains 200 words where the term "kill" is found three times.

TF(t)= (3 / 200) = 0.015.

IDF(t) = log(1000 / 10) = 2.

The Tf-idf weight = 0.015 * 2 = 0.03.

It can be noted that the texts of some authors are consistent, or at least similar, in the frequency of use of different words. The current work focused on word frequency as Tweets are limited in length (280 characters). The analysis of text was carried out as follows (for instance, for "tf-idf uni-gram" where the word count is 3315):

- The first step is that a vector of equivalent size as the word count (3317) is made. Then for every position in the vector a word is associated with 0.

word1	word2		Word3317
0	0	……….	0

- When a tweet is parsed and a word is found in the tweet.

- First TF value for that word is calcuated.

- Then the IDF value for that word is calulated.

- Then finally both values are multiplied to get the TF-IDF weight value.

- Then this value is placed in the matrix for that word.

- For each tweet one single row in the matrix is assigned.

- Steps are repeated for every different word in the tweet.

- In this way the matrix is created with TF-IDF weights.

An example of building a word vector can be considered on the following tweet:

لجيش_ا ينتصر"

Translation:” Islam_army will prevail”

Suppose the

Total number of tweets = 346 and Islam_army is in 35 tweets

TF(“Islam_army”) = 1/3 = 0.3333….

IDF(“Islam_army”) = log(350/35) = 0.99500805444

Thus TF-IDF = 0.3333 x 0.99500805444 = 0.3316

Now suppose

Total number of tweets = 346 and prevail is in 50 tweets

TF(“Islam_army”) = 1/3 = 0.3333

IDF(“Islam_army”) = log(346/50) = 0.84010609445

Thus TF-IDF = 0.3333 x 0. 84010609445 = 0.28000736128

Word 1	Islam_army	….	prevail	…	Word 3317	Word 3317
0	0.3316	0	0.2800	0	0	0

At the end, the sum of all values contained in the vector will always be 1. Hashtags are used to emphasise the meaning and importance of some words and to permit users to easily find messages with specific themes or content. The Hashtags were compiled based on the literature review (Ali, 2016; Omer, 2015; Salmi Abdellatif, 2016), with the avoidance of hashtags that do not exist in Arabic or they are inactive for more than a month. The entire list of hashtags can be found in the Appendix.

The vector space model was used to represent the tweets and introduce the use of syntactic n-grams as markers of personality along with the use of ML classifiers. After creating the vectors, there are three vectors (tf-idf uni-gram, tf-idf bi-gram, tf-idf ti-gram) and the data evaluated by the experts are input to the machine learning techniques (five workbooks) to give the output. Black-box models, such as support vector machines or artificial neural networks, do not allow such an interpretation, and can only be verified externally (Dreiseitl, & Ohno-Machado, 2002; Wolpert & Macready, 1997).

3.4 Training and Testing Model

This part discusses Creating Arabic terrorism model using Machine Learning algorithms and text classification.

3.4.1 Classification Algorithms

The response in our experiment is qualitative. Predicting a qualitative response for an observation can be referred to as classifying that observation, since it involves assigning the observation to a category, or class. There are many possible classification techniques, or classifiers, that one might use to predict a qualitative response. In the classification setting, there is a set of training observations (x1, y1), . . . , (xn, yn) that can be used to train a classifier. In this work, the observation is textual data. Also, in the dataset in this study, there are labels for all the observations. So, the experiment with a few supervised text classification methods have been explained in the following section. The data set is separated into two categories (training and testing). Thus supervised learning approaches such as NB, SVM, Logistic Regression, and 2 AdaBoost (discrete, real) were implemented to extract the terrorism detection. Finally, the results were evaluated through matrix confusion that is used to calculate the proposed benchmark accuracy.

3.4.2 Text Classification

Naïve Bayes is a classification method based on Bayes' theorem that derives the probability of the given feature vector being associated with a label

Naive Bayes is a grouping technique linked to the tBayes that connotes the possibility of the provided feature vector as linked to a label (Maron, 1961). Naive Bayes has a primitive presumption of a situational freedom for every provided feature. These points to the fact that features are expected to be independent by algorithm and this is not always like that (Maron, 1961). It appears like a model that is generative, meaning that joints are modeled by it, especially the distribution ones, that deals with target Y and feature X. What follows is the prediction of posterior possibility as provided as P(X/Y). The squash value and the linear function are taken as output in Logistic regression within a given range of (0, 1) and this is when logistic function (sigmoid function) is applied (Walker and Duncan, 1967). The sigmoid function is s-shaped as displayed in figure 2, and any valued number that is real can go with it and can be placed anywhere between the intervals of 0 and 1, and not strictly to those intervals. Label 0 or 1 is assigned when the squash value appears typically greater than a threshold value. This gives justification to the name 'logistic regression'. It maximizes the possibility of log that a data is pointed at random and correctly gets classified. Mathematically, it is referred to as MLE (Maximum Likelihood Estimation).

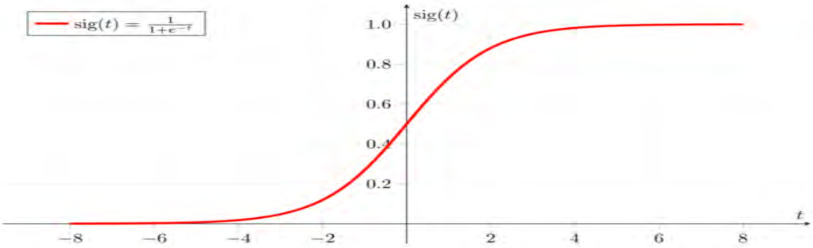

Figure 3.2: Sigmoid curve

The objective of the support vector machine algorithm is to find the hyperplane that has the maximum margin in an N-dimensional space (N—the number of features) that distinctly classifies the data points (Cortes and Vapnik, 1995). Data points falling on either side of the hyperplane can be attributed to different classes. Also, the dimension of the hyperplane depends upon the number of features. If the number of input features is n, then the hyperplane is of dimension n-1. Support vectors are data points that are closer to the hyperplane and influence the position and orientation of the hyperplane. Using these support vectors, the margin of the classifier is maximized.

Boosting is a general approach that can be applied to many statistical learning methods for regression or classification. It is particularly popular in the context of decision tree. Boosting works by fitting a separate decision tree in each iteration which is grown sequentially using information from previously grown trees. Boosting does not involve bootstrap sampling like bagging. Instead each tree is fit on a differently weighted version of the original data set. Adaboost (Kégl , 20 December 2013) is one of the popular boosting algorithm in which the model learns slowly. Given the current model, a decision tree is fitted to the residuals from the model. That is, a tree using the current residuals (Y- Y') is fitted, rather than the outcome Y, as the response. Then this new decision tree is added into the fitted function in order to update the residuals. Each of these trees can be rather small, with just a few terminal nodes. By fitting small trees to the residuals, the model is slowly improved in areas where it does not perform well. The AdaBoost algorithm is known as "**Discrete AdaBoost**" if the base classifier returns a discrete class label. If the base classifier instead returns a real-valued prediction (e.g., a probability mapped to the interval [−1, 1]), the algorithm is called "**Real AdaBoost**".

Naïve Bayes is the simplest of all the algorithms and naturally it is the fastest of all the algorithms. Even though it is the simplest algorithm, it performs reasonably well with text classification and often is a good baseline or starting point. **Logistic Regression(LR)** produces probabilistic values while **SVM** produces 1 or 0. So **LR** does not make absolute prediction and it does not assume data is enough to give a final decision. This maybe be a good property when what wanted is an estimation or there is no high confidence in the data. But LR is not perfect in every sense. **LR** minimizes logistic loss to classify data points. Logistic loss diverges faster than hinge loss of SVM. So, in general, it will be more sensitive to outliers. So to make sure our algorithm is not affected by outliers, using **SVM** is a good choice. Also **SVM** is highly robust and it works with not linearly separable data where **LR** fails. **Adaboost** is a relatively complex ensemble algorithm which uses multiple learners and usually less prone to overfitting and provide better classification performance than other algorithms such as **LR** or **SVM**. So it's always worth trying **Adaboost** to see whether it is able to provide additional benefits.

In text classification a tweet is classified into a predefined category. The category is a Boolean value indicating if the tweet contains radical content or not. After various text classification algorithms were discussed, the study arrived at the general text classification process which includes the following steps:

- preprocessing the text
- creating feature vectors
- selecting features

- Training a model on train data.
- Calculating classification accuracy on test data.

3.5 Evaluation Measures

In order to decide whether the classifiers are accurately capturing a pattern, the model will be evaluated. Validation is an important step that allows testing the accuracy of the classifiers. Therefore, the result of this evaluation is crucial for determining the trustworthiness of the classifiers. One method for evaluating the performance of the classifiers is the cross-validation method. The cross validation method splits the data into two parts: testing data and training data. Moreover, cross-validation performs multiple evaluations on different test sets, and combines the scores from these evaluations. In particular, the k-fold cross-validation method was used, in which the data was divided into k parts. The k-fold cross-validation form is considered to be the basic form of cross-validation. Several methods were used for evaluating the proposed classifiers' performance and comparing them with each other:

- Confusion matrices
- True of' values (positive TP, negative TN)
- False of' values (positive FP, negative FN).
- Precision and Recall
- F-measure and Accuracy
- Confusion Matrices

A confusion matrix, also known as an error matrix, is a table that contains information on the actual classification performance of a classification system. Each column of the table represents an object in a predicted class, while each row represents the object in an actual class.

- Accuracy

Accuracy simply measures the percentage of inputs in the test set that are correctly classified. The accuracy of the classifier was calculated as follows:

$$accuracy = \frac{\sum true\, positive + \sum true\, negative}{\sum total\, population}$$

- Precision

Precision is defined as the percentage of the predicted sentence class that is correctly classified. The precision for the positive class, for instance, is calculated as follows:

$$\text{Precision} = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

- Recall:

Recall is a measure of completeness; it calculates the percentage of the total documents for the given class that are correctly classified. Recall is defined as the probability of a random sentence to be classified within this class. Recall for the positive class, for instance, is calculated as follows:

$$\text{Recall} = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

- F-Measure:

F-Measure can be defined as a harmonic average of the precision and the recall values obtained. The F-measure gives a good indication of the overall performance of a classifier and is calculated using the following formula:

$$\text{F1} = 2 \text{ x } \frac{Precision * Recall}{Precision + Recall}$$

3.6 Summary

This chapter explained the different phases of work undertaken in this study. The study was divided into three main phases, namely, data analysis, feature extraction, and training model which were discussed in detail in this chapter.

CHAPTER 4: RESULT AND EVALUATION

4.1 Chapter Overview

This chapter discusses the findings of this study. In addition, it elaborates on the methods used to evaluate the performance of classification models: precision, recall, F-measure, and accuracy.

4.2 Experimental Setup

In this chapter, the dataset was constructed by collecting tweets from tweeter, and after collecting the tweets, validity and reliability statistical measurements were applied to standardize and evaluate the dataset.

Table 4.1 : Experimental Setup of Terror orientation detection

Name	Description
Data collecting	Twitter
Validity	Statistical approach
Number of tweet	346
Language programing	Python 3.6
NLP tools	NLTK, API python
Classification tools	Sklearn
Classification Methods	-SVM - Naïve Basie - Logistic Regression - AdB_SAMME -AdB_SAMME.R
Statistical Programs	SPSS version 24

Table 4.2 shows the following expermintal setup used in this research:

- Step1 The dataset was collected from the twitter using python API techniques.
- Step2 The statistical analysis approaches were applied to evaluate the collected dataset in the step (1) using SPSS version 24.
- Step3 The feature extraction approaches of text were vectorized using several approaches discussed in section 3.3.2.
- Step4 The feature extraction vectors was trained and tested by several classification approach showed in table 4.2 using the Sklearn library in python.

4.3 Dataset and Statistical Analysis

The Data was collected using Twitter API and was further built as a data set. Twitter search was used to search for tweets using the Arabic language. The data were collected from Twitter during the month of March of 2018. The focus was on tweets in the Arab area through the identification of the center, which is Riyadh city in Saudi Arabia and the diameter of the circle is 2000 km. Each extracted tweet contained extensive information, such as user ID, tweet text, and number of tweets of a user. The number of collected tweets was 135,069. However, the number of tweets was reduced to 346 tweets.

After the collection of the tweets, the data were preprocessed. These tweets were given to human experts in the form of a questionnaire, and each point was a tweet and there was a questionnaire containing 346 sentences. There were two options for each sentence and the options were whether each sentence supported terrorism or did not support terrorism. In other words, whether the sentence was supportive of terrorism or against terrorism.

In this work, the collected data was divided into two datasets that were presented in the form of a questionnaire for four arbitrators from the Ministry of Education in Saudi Arabia: three teachers and an educational supervisor who is attached to the educational administration of the ministry. They are native speakers of Arabic and are well acquainted with the topic.

The questionnaire consisted of 346 tweets, and each tweet, contained two responses. Each expert determined whether the sentence supported terrorism or not, for all sentences in the questionnaire. The four arbitrators judged a sample of 346 tweets and for each condition in the tweet the arbitrators judged as to whether it was positive or negative. There was agreement on all tweets except seven tweets.

A number of words were adopted based on the research related to the search for terrorist groups and some common terms they used in their tweets. Duplicate tweets and tweets sent from news accounts were excluded. Tweets written in a language other than Arabic and unrelated tweets were also excluded.

Table 4.2 : Cronbach's Alpha

Reliability Statistics	
Cronbach's Alpha	No. of Judges
0.994	4

The tweets were given to experts who evaluated and judged the tweets as to whether they were radical or not radical. The validity and reliability of the dataset were measured statistically using Cronbach's Alpha as shown in table 4.1. As is statistically known, the value of the Cronbach's Alpha must be more than 0.8 to judge the validity and reliability of the data. Therefore, this work can use the data and conduct experiments on them.

4.4 Experiment Result Analysis

In this work different classification techniques, datasets, and features were used to evaluate the effectiveness of the techniques used in this experiment to determine extremist tweets. Most of the features that are important to make this task easy were examined and a total of 346 tweets were used. Each tweet was classified into a predefined category to indicate if the tweet supported terrorism or not. Two measures were used to evaluate the precision of the classification of the tweets, namely, F-measure and Accuracy. F1 Score might be a better measure to use if seeking a balance between recall and precision on uneven class distribution (large number of Actual Negatives).

4.5 The Experiment Analysis using F1-Score

Five classfiers (AdB_SAMME, AdB_SAMME.R, Linear, NB and LR) were applied on each of the 3 features (S1, S2 and S3) separately. So a total of 15 experiments were conducted. In this section, class wise and average Precision, Recall and F-measure for all 15 experiments are presented. Precision helps to understand what percent of predictions made by the model is correct. It is mathematically defined by TP/(TP + FP). Recall helps in understanding what percent of the positive cases is the model able to predict. It is mathemically definted by TP/(TP+FN).

[TN / True Negative: label was negative and the model predicted negative

TP / True Positive: label was positive and the model predicted positive

FN / False Negative: label was positive but the model predicted negative

FP / False Positive: label was negative but the model predicted positive]

In the tables below, “Support” column represents the total number of samples from each class. In “avg/total” row, “avg” refers to the first 3 columns, i.e “precision”, “recall” and “f1-score”, whereas “total” refers only to “Support”. “avg” is calculatd by summing up weighted value of each class. For example, how the “avg/total” row is calculated is presented in table 4.3.

“avg” for precision = .72*(146/(146+200))+.92*(200/(146/200)) ≈ .84

“avg” for recall = .91*(146/(146+200))+.74*(200/(146/200)) ≈ .82

“avg” for f1-score = .81*(146/(146+200))+.82*(200/(146/200)) ≈ .82

“total” for support = (146+200) = 346

F-measure is also called the F-Score. In other words, the F1 score conveys the balance between precision and the recall.

Table 4.3: Feature extraction of unigram sample by AdB_SAMME classification model

	precision	recall	f1-score	Support
support	0.72	0.91	0.81	146
not support	0.92	0.74	0.82	200
avg / total	0.84	0.82	0.82	346

Table 4.4: Feature extraction of S1 by AdB_SAMME.R classification model

	precision	recall	f1-score	Support
support	0.72	0.91	0.81	146
not support	0.92	0.74	0.82	200
avg / total	0.84	0.82	0.82	346

Table 4.5: Feature extraction of S1 by Linear classification model

	precision	recall	f1-score	Support
support	1.00	0.99	0.99	146
not support	0.99	1.00	1.00	200
avg / total	0.99	0.99	0.99	346

Table 4.6: Feature extraction of S1 by NB classification model

	precision	recall	f1-score	Support
support	0.99	1.00	0.99	146
not support	1.00	0.99	0.99	200
avg / total	0.99	0.99	0.99	346

Table 4.7: Feature extraction of S1 by LR classification model

	precision	recall	f1-score	Support
support	1.00	0.95	0.98	146
not support	0.97	1.00	0.98	200
avg / total	0.98	0.98	0.98	346

Table 4.8: Feature extraction of : S2 by AdB_SAMME classification model

	Precision	recall	f1-score	Support
support	0.72	0.91	0.81	146
not support	0.92	0.74	0.82	200
avg / total	0.84	0.82	0.82	346

Table 4.9: Feature extraction of S2 by AdB_SAMME.R classification model

	precision	recall	f1-score	Support
support	0.72	0.91	0.81	146
not support	0.92	0.74	0.82	200
avg / total	0.84	0.82	0.82	346

Table 4.10: Feature extraction of S2 by Linear classification model

	Precision	recall	f1-score	Support
support	1.00	0.99	1.00	146
not support	1.00	1.00	1.00	200
avg / total	1.00	1.00	1.00	346

Table 4.11: Feature extraction of S2 by NB classification model

	precision	recall	f1-score	Support
support	0.99	1.00	0.99	146
not support	1.00	0.99	0.99	200
avg / total	0.99	0.99	0.99	346

Table 4.12: Feature extraction of S2 by LR classification model

	precision	recall	f1-score	Support
support	1.00	0.98	0.99	146
not support	0.99	1.00	0.99	200
avg / total	0.99	0.99	0.99	346

Table 4.13: Feature extraction of S3 by AdB_SAMME classification model

	precision	recall	f1-score	Support
support	0.73	0.91	0.81	146
not support	0.92	0.75	0.83	200
avg / total	0.84	0.82	0.82	346

Table 4.14: Feature extraction of S3 by AdB_SAMME.R classification model

	precision	recall	f1-score	Support
support	0.88	0.73	0.79	146
not support	0.82	0.93	0.87	200
avg / total	0.84	0.84	0.84	346

Table 4.15: Feature extraction of S3 by Linear classification model

	precision	recall	f1-score	Support
support	1.00	0.99	1.00	146
not support	1.00	1.00	1.00	200
avg / total	1.00	1.00	1.00	346

Table 4.16: Feature extraction of S3 by NB classification model

	precision	recall	f1-score	Support
support	0.99	1.00	0.99	146
not support	1.00	0.99	0.99	200
avg / total	0.99	0.99	0.99	346

Table 4.17: Feature extraction of S3 by LR classification model

	precision	recall	f1-score	Support
Support	1.00	0.98	0.99	146
not support	0.99	1.00	0.99	200
avg / total	0.99	0.99	0.99	346

Tables 4.3 to 4.17 show the details of the precision, recall and F-score. There is a disparity in the values of classification based on the quality of the classification technique, and thus it is generally apparent that there is an improvement in reading the tendencies of the tweets. Besides, as shown in the tables, the linear classification technique achieved the best value of the F-score due to the reason that the polarity has only two directions (terrorism or not-terrorism).

4.6 Discussion

In this section, the results between the classification approaches are discussed. Table 4.18 shows that the accuracy of all the samples and the results differed depending on what datasets were used and also the features (S1 + S2 + S3). As can be noted, linear SVM performs very well with 99.7 % classification accuracy on the classification. S2 features (bigram) seem to be the most effective in this case. However, if the other classification accuracy is looked at, it can be seen that S3 (trigram) features tend to perform slightly better than unigram and bigram features. The fact that Linear SVM, NB, and LR largely outperform Adaboost may suggest that the data distribution is linear and simple enough that complex algorithms like adaboost is not really necessary here.

Table 4.18: The results for technique and the features (S1+S2+S3)

Sample	Technique	Accuracy	Accuracy
S1	AdB_SAMME	0.815028901734104	0.815
S1	AdB_SAMME.R	0.815028901734104	0.815
S1	Linear (SVM)	0.9942196531791907	0.994
S1	NB	0.9942196531791907	0.994
S1	LR	0.9797687861271677	0.980
S2	AdB_SAMME	0.815028901734104	0.815
S2	AdB_SAMME.R	0.815028901734104	0.815
S2	Linear (SVM)	0.9971098265895953	0.997
S2	NB	0.9942196531791907	0.994
S2	LR	0.9913294797687862	0.991
S3	AdB_SAMME	0.8179190751445087	0.820
S3	AdB_SAMME.R	0.8410404624277457	0.841
S3	Linear (SVM)	0.9971098265895953	0.997
S3	NB	0.9942196531791907	0.994
S3	LR	0.9913294797687862	0.991

Figure 4.1 shows the results for five different classifiers using all features on all the datasets. As can be seen in the table 4.18, Linear performs slightly better than other classifiers. The accuracy, when using the Linear classifier, is still high (99.7%).

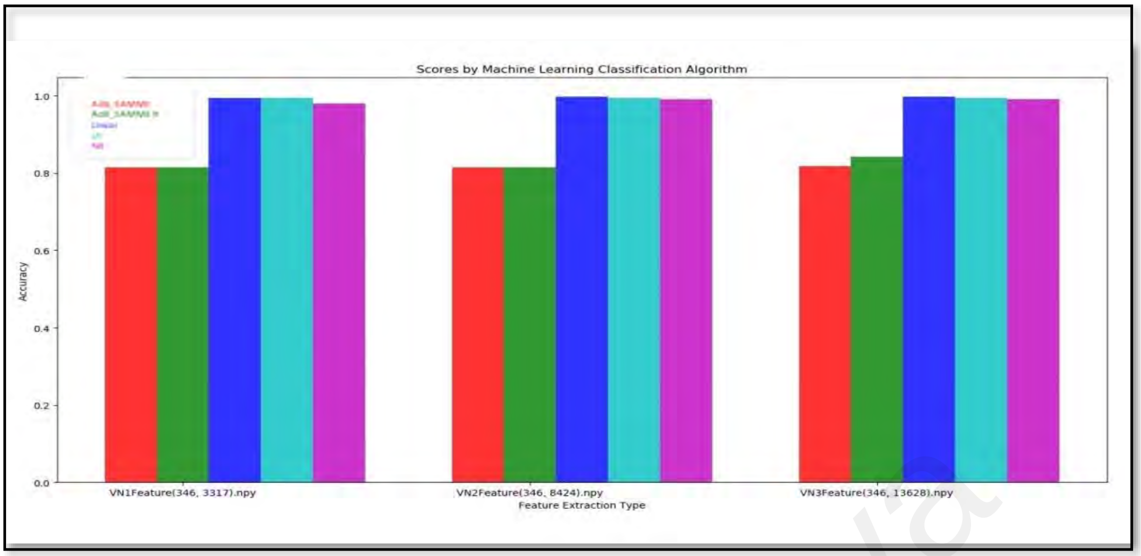

Figure 4.1: Scores by Machine Learning Classification Algorithms

In addition to the classification performance, the training time is a second important factor that affects the suitability of a classification algorithm. An algorithm with a slightly lower accuracy maybe preferred if its training time is significantly lower, especially if the time factor is important. The tables (4.19 to 4.21) show the training times for each classifier.

In the following table, “tweet_no” means “total number of tweets in the dataset”, “FE_no” means “number of features”. [for unigram, the “FE_no” signifies number of unique words in the entire dataset. for bi-gram, the “FE_no” signifies number of unique consecutive word pair. for tri-gram, the “FE_no” signifies the number of unique consecutive word triplets.

“ML_Name” signifies the classifier used for classification task.

“accuracy”, “f1_score” and “time” are self-explanatory.

Ideally, accuracy is needed and f1_score need to be high and time need to be as low as possible. So if both classification performance and time are considered, it can be seen that the 4th experiment(uni-gram features with NB Classifier) ranks higher than Linear Classifier.

Table 4.19: The results for the performance time (tf-idf uni-gram)

no	tweet_no	FE_no	ML_name	Accuracy	f1_score	time
1	346	3317	AdB_SAM	81.50289	81.59703	0.207
2	346	3317	AdB_SAM	81.50289	81.59703	0.139
3	346	3317	Linear	99.42197	99.42141	0.808
4	346	3317	NB	99.42197	99.42248	0.05
5	346	3317	LR	97.97688	97.96943	0.004

Table 4.20: The results for the performance time (tf-idf bi-gram)

no	tweet_no	FE_no	ML_name	Accuracy	f1_score	time
1	346	8424	AdB_SAM	81.50289	81.59703	0.295
2	346	8424	AdB_SAM	81.50289	81.59703	0.561
3	346	8424	Linear	99.71098	99.71085	2.113
4	346	8424	NB	99.42197	99.42248	0.13
5	346	8424	LR	99.13295	99.13167	0.015

Table 4.21: The results for the performance time (tf-idf ti-gram)

no	tweet_no	FE_no	ML_name	Accuracy	f1_score	time
1	346	13628	AdB_SAM	81.79191	81.88781	0.531
2	346	13628	AdB_SAM	84.10405	83.82738	0.549
3	346	13628	Linear	99.71098	99.71085	3.428
4	346	13628	NB	99.42197	99.42248	0.214
5	346	13628	LR	99.13295	99.13167	0.018

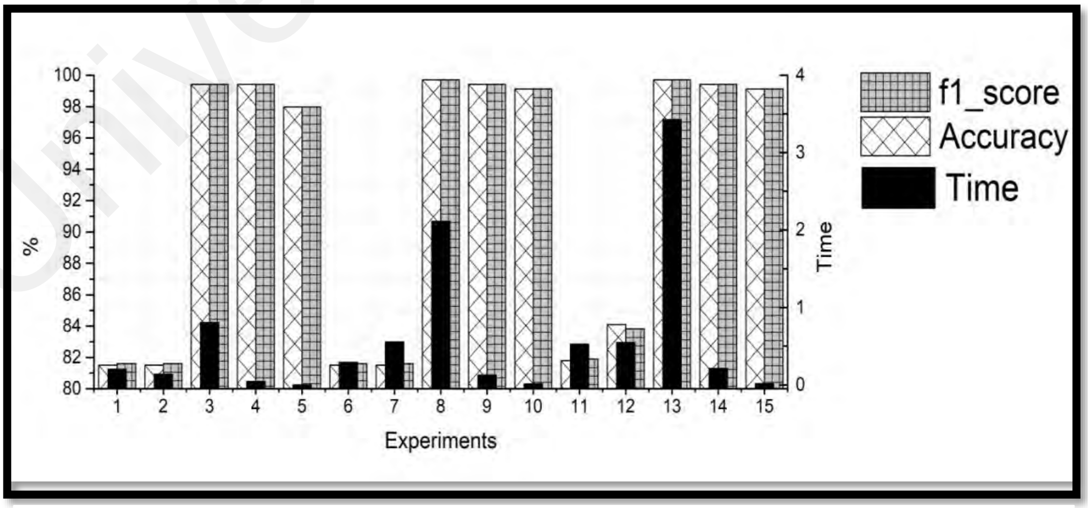

Figure 4.2 : The time and accuracy of classification performance

The following table shows the similarities and differences in accuracy compared to previous studies, taking into consideration the different languages (English, Arabic):

Table 4.22: Similarities and differences in accuracy

	Omer (2015)	FE	My approach	FE	Delta ∆
AdB_SAMME.R	100	(text,Time)	84.1	(text)	-0.159
NB	89.0	(text,Time)	99.4	(text)	0.117
SVM	97.9	(text,Time)	99.7	(text)	0.018
LR	-		99.1	(text)	-
AdB_SAMME	-		81.8	(text)	-

Although, there are several datasets in the Arabic language in twitter, these datasets are in the sentiment analysis approaches and there is research on terrorism detection in the Arabic language is rare. Table 4.23 shows the comparision of results in twitter according to the accuracy of sentiment analysis.

Table 4.23: Accuracy comparison between approaches

Research	Domain	ML	Accuracy
Deng(2014)	Opinion	SVM	75.90%
Rui et al.(2014)	Social comment	SVM	76.78%
Li and Li (2017)	Social news	SVM	83.80%
Farra et al. (2017)	Reviews	SVM	87.43%
Ali, F., Khan (2018)	Terrorsim English	SVM	87.90%
My Proposal	Terrorsim	SVM	89.24%

The accuracy in Table 4.23 shows the Arabic terrorism detection based on the dataset collected for this research compared of the other Arabic datasets that are used in the opinion sentiment analysis. In addition, the Ali dataset that was used in the terrorism in the English language achieved 87.90% in the terrorism detection using the same feature extraction and classification approaches. The difference between the approach and the English approach is due to the fact that this study's dataset is rich in positive terrorism tweets more than the English tweets. There is another reason that enhances the accuracy in this study and that is, the location for the current study is more significant than the one that was used in the English dataset.

4.7 Summary

In this chapter, the details of the implementation phase and evaluation analysis were discussed. The overall analysis highlighting the accuracy, precision, recall and F-measure were also provided.

CHAPTER 5: CONCLUSION AND FUTURE STUDIES

5.1 Chapter Overview

This chapter draws the general research conclusion and also discusses the future directions for this research. Additionally, the contribution of this research is outlined.

In this study, the interest was in the issue of detecting radical content (in Arabic) on the social media, and in particular, on Twitter.

5.2 Conclusion

In this work an approach to classify Arabic tweets as supporting or not supporting radical content was presented. There are many types of machine learning algorithms that differ in their approaches. Supervised learning was used in the current study. The study included extraction of textual features. The results of the experiments proved that the highest accuracy was achieved by the linear classifier. Different datasets were used in the present study. To build classifiers for the present study tweets that supported radical groups and tweets having messages oriented against ISIS were collected .

Experiments were run using different text classifiers such as AdaBoost, Naive Bayes, SVM and LR. The results that were obtained suggest that classification is a feasible way of identifying radical content online in the Arabic language, and, in particular, on the social media platform, Twitter. In the future, there is a need to replicate these results with even more diverse and/or complex datasets.

The research goal was to develop a Machine Learning Approach to detect the terrorism tweets in Twitters. In this thesis, an improvement was made on the initial new dataset of Arabic terrorism messages that were extracted by API from Twitter in the Arabic language. The proposed dataset was evaluated by the statistical rules and tools with significant validity and reliability. In addition, the approach investigated and improved the predictions of the radical tweets in twitters.

It has been proven that the proposed model outperformed in extracting the polarity of the radical content using several language models (n-gram) and co-occurrence vector approaches such as count vector, term frequent, and term frequent inverse document frequent.

The results of the investigation demonstrated the ability of the proposed approach to extract and select the features from the tweets. The experimental results have shown that the proposed feature selection and extraction based on the machine learning approaches improved the detection of the radical polarity of tweets by considering the vector of features trained by machine learning.

The proposed model has been evaluated using performance metrics such as precision, recall, F-score and accuracy and statistical measurement in validity and reliability. The results showed that the proposed techniques can enhance tweets polarity in the radical content.

Finally,this study tried to answer the research questions. Initially, the most important features through which we can find out the direction of the text and its terrorist tendencies were identified and identified. Several machine learning models were applied to these features to extract their internal implications and insights. After obtaining the outputs of machine learning was evaluated through a set of measurement tools recognized in this area. The output results are encouraging to use proposed approach for Arab terrorist tweets detection.

5.3 Contributions

In summary, a viable way to detect radical content (in Arabic) on the social media, and, in particular, on Twitter was proposed to extract the radical polarity of tweets. Further, the main contributions of this research are as follows:

- Firstly, implemention a set of features and running a set of experiments to select the best machine learning : This study used a machine learning model that automatically detected Arabic tweets from terrorist groups on the Twitter platform. This work has investigated the use of five text classifiers , due to the variety of philosophies behind each method and its learning process, to select the best classifier for the features in this study. Then, the model assists analysts in their working with detecting Arabic radical content on Twitter and helps them to find the various extremist messages sent by terrorists, particularly in the Arabic world.
- Secondly, an effective model was built for terrorism detection : the final outcomes of this research is a machine learning model that can detect radical Arabic content on social media platforms that has a great potential for yielding results.
- Finally, evaluating the machine learning approaches that were aiding to take the necessary procedures and suspending those active accounts in support of terrorism.

5.4 Limitations

Although the algorithms applied in this study contributed towards detection and categorization of a set of diverse Arabic tweets, and finally, classified them as either radical or not radical, there are some limitations. The main limitation in this work is the size of the dataset. In the future this study should be expanded and a bigger dataset should be used to confirm and improve the results. In addition, there are several drawbacks in using a language model in the feature extraction such as the redundancy and huge feature size.

5.5 Future Studies

One way to improve this study is to classify a Twitter account as being radical or not. In addition, the number of features should be reduced and important features should be selected that will overcome the complexity in the training model. The promising results obtained in this study provides optimism to expand the area of collection to include different tweets around the world, and this leads to further work and research on Arabic tweets and the extraction of features. Also, there is suggestion to applay the embedding approach to extract new features related to terrorism field in Arabic languge.

REFERENCES

Agarwal, S., & Sureka, A. (2015, February). Using known and SVM based one-class classifier for detecting online radicalization on twitter. In *International Conference on Distributed Computing and Internet Technology* (pp. 431-442). Springer, Cham.

Aistrope, T. (2016). Social media and counterterrorism strategy. *Australian Journal of International Affairs*, *70*(2), 121-138.

Al-Arman Mohammed Saad, & Al-Shawabkeh Mohammed Abdullah (2014). To correct the concept of terrorism, violence, vandalism and ignorance between Sharia and law. *Journal of the researcher for academic studies. 1 (3), 28-59.*

Al-Qatab, B. A., & Ainon, R. N. (2010, June). Arabic speech recognition using hidden Markov model toolkit (HTK). In *Information Technology (ITSim), 2010 International Symposium in* (Vol. 2, pp. 557-562). IEEE.

Al-Mazari, A., Anjariny, A. H., Habib, S. A., & Nyakwende, E. (2018). Cyber terrorism taxonomies: Definition, targets, patterns, risk factors, and mitigation strategies. *International Journal of Cyber Warfare and Terrorism*, *6*(1), 1-12.

Ali, G. A. (2016). Identifying Terrorist Affiliations through Social Network Analysis Using Data Mining Techniques.

Ali, F., Khan, F. H., Bashir, S., & Ahmad, U. (2018, October). Counter Terrorism on Online Social Networks Using Web Mining Techniques. *In International Conference on Intelligent Technologies and Applications (pp. 240-250). Springer, Singapore.*

Almas, Y., & Ahmad, K. (2007, July). A note on extracting 'sentiments' in financial news in English, Arabic & Urdu. In *The Second Workshop on Computational Approaches to Arabic Script-based Languages* (pp. 1-12).

Alsaedi, N. (2017). *Event Identification in Social Media using Classification-Clustering Framework* (Doctoral dissertation, Cardiff University). Retrieved from https://orca.cf.ac.uk/100998/1/2017alsaedinphd.pdf

Alshari, E. M., Azman, A., Doraisamy, S., Mustapha, N., & Alkeshr, M. (2017, August). *Improvement of Sentiment Analysis Based on Clustering of Word2Vec Features.* In *28th International Workshop on Database and Expert Systems Applications* (DEXA), held on 29 August 2017 (pp. 123-126). IEEE. Retrieved from https://www.uni-weimar.de/medien/webis/events/tir-17/tir17-talks/Eissa2017_improvement-of-sentiment-analysis-based-on-clustering-of-word2vec-featuers_presentation.pdf

Amplayo, R. K., & Occidental, J. (2015). Multi-level classifier for the detection of insults in social media. In *Proceedings of 15th Philippine Computing Science Congress*.

Anjaria, M., & Guddeti, R. M. R. (2014, January). Influence factor based opinion mining of Twitter data using supervised learning. *In 2014 Sixth International*

Conference on Communication Systems and Networks (COMSNETS) (pp. 1-8). IEEE.

Aramaki, E., Maskawa, S., & Morita, M. (2011, July). Twitter catches the flu: detecting influenza epidemics using Twitter. *In Proceedings of the conference on empirical methods in natural language processing (pp. 1568-1576). Association for Computational Linguistics.*

Auria, L., & Moro, R. A. (2008). *Support vector machines (SVM) as a technique for solvency analysis*. DIW Berlin German Institute for Economic Research.

Ashcroft, M., Fisher, A., Kaati, L., Omer, E., & Prucha, N. (2015, September). Detecting jihadist messages on twitter. In *Intelligence and Security Informatics Conference, 2015 European* (pp. 161-164). IEEE.

Atefeh, F., & Khreich, W. (2015). *A survey of techniques for event detection in twitter. Computational Intelligence*, *31*(1), 132-164.

Badjatiya, P., Gupta, S., Gupta, M., & Varma, V. (2017, April). Deep learning for hate speech detection in tweets. *In Proceedings of the 26th International Conference on World Wide Web Companion (pp. 759-760).* International World Wide Web Conferences Steering Committee.

Benhardus, J., & Kalita, J. (2013). *Streaming trend detection in twitter. International Journal of Web Based Communities, 9(1), 122-139.*

Beninati, J. A. (2016). *Examining the cyber operations of ISIS* (Doctoral dissertation, Utica College).

Bernatis, V. (2014). *The Taliban and Twitter: Tactical reporting and strategic messaging. Perspectives on Terrorism*, *8*(6), 25-35

Bird, G., Blomberg, S. B., & Hess, G. D. (2008). International terrorism: Causes, consequences and cures. *World Economy*, *31*(2), 255-274.

Bjoergum, M. C. H. (2014). *The Credibility of News Media: The difference in framing between traditional media and Twitter after the Boston Marathon bombing* (Doctoral dissertation, Hawaii Pacific University).

Blaker, L. (2015). The Islamic State's use of online social media. *Military Cyber Affairs, 1*(4), 1 – 9.

Blomberg, S. B., & Hess, G. D. (2008). From (no) butter to guns? Understanding the economic role in transnational terrorism. *Terrorism, economic development, and political openness*, 83-115.

Bodine-Baron, E., Helmus, T. C., Magnuson, M., & Winkelman, Z. (2016). *Examining ISIS Support and Opposition Networks on Twitter.* RAND Corporation Santa Monica United States.

Bowyer, C. E. (2015). *Twitter and the Islamic State: What is the Government's Role?* (MA Thesis, American Public University: Charles Town).

Brahmi, A., Ech-Cherif, A., & Benyettou, A. (2012). Arabic texts analysis for topic modeling evaluation. *Information Retrieval, 15*, 33–53.

Burnap, P., Williams, M. L., Sloan, L., Rana, O., Housley, W., Edwards, A., & Voss, A. (2014). *Tweeting the terror: modeling the social media reaction to the Woolwich terrorist attack. Social Network Analysis and Mining*, *4*(1), 206 – 220.

Caplan, J. (2013). Social media and politics: Twitter use in the second congressional district of virginia. *Elon Journal of Undergraduate Research in Communications , 4(1).*

Carchiolo, V., Longheu, A., & Malgeri, M. (2015, September). Using twitter data and sentiment analysis to study diseases dynamics. *In International Conference on Information Technology in Bio-and Medical Informatics (pp. 16-24). Springer, Cham.*

Cheong, M., & Lee, V. C. (2011). A microblogging-based approach to terrorism informatics: Exploration and chronicling civilian sentiment and response to terrorism events via Twitter. *Information Systems Frontiers*, *13*(1), 45-59.

Chae, B. K. (2015). Insights from hashtag# supply chain and Twitter Analytics: Considering Twitter and Twitter data for supply chain practice and research. *International Journal of Production Economics*, *165*, 247-259.

Chatfield, A. T., Reddick, C. G., & Brajawidagda, U. (2015, May). Tweeting propaganda, radicalization and recruitment: Islamic state supporters multi-sided twitter networks. In *Proceedings of the 16th Annual International Conference on Digital Government Research* (pp. 239-249). ACM.

Chen, G., Li, X., Chen, J., Zhang, Y. N., & Peijnenburg, W. J. (2014). Comparative study of biodegradability prediction of chemicals using decision trees, functional trees, and logistic regression. *Environmental toxicology and chemistry*, *33*(12), 2688-2693.

Chin, D., Zappone, A., & Zhao, J. (2016). Analyzing Twitter sentiment of the 2016 presidential candidates. *American Journal Of Science and Research.*

Choi, D., Ko, B., Kim, H., & Kim, P. (2014). Text analysis for detecting terrorism-related articles on the web. *Journal of Network and Computer Applications, 38*, 16-21.

Choi, D., Hwang, M., Ko, B., & Kim, P. (2011). Solving English questions through applying collective intelligence. In *Future Information Technology* (pp. 37-46). Springer, Berlin, Heidelberg.

Choi, D., Ko, B., Lee, E., Hwang, M., & Kim, P. (2012, September). Automatic evaluation of document classification using n-gram statistics. In *2012 15th International Conference on Network-Based Information Systems* (pp. 739-742). IEEE.

Choudhary, P., & Singh, U. (2015). A survey on social network analysis for counter-terrorism. *International Journal of Computer Applications*, *112*(9), 24 – 29.

Cohen, K., Johansson, F., Kaati, L., & Mork, J. C. (2014). Detecting linguistic markers for radical violence in social media. *Terrorism and Political Violence, 26(1), 246-256.*

Collins, M., Schapire, R. E., & Singer, Y. (2002). Logistic regression, AdaBoost and Bregman distances. *Machine Learning, 48*(1-3), 253-285.

Corner, E., Gill, P., & Mason, O. (2016). Mental health disorders and the terrorist: A research note probing selection effects and disorder prevalence. *Studies in Conflict & Terrorism, 39*(6), 560-568.

Cortes, Corinna; Vapnik, Vladimir N. (1995). "Support-vector networks". Machine Learning. 20 (3): 273–297. CiteSeerX 10.1.1.15.9362. doi:10.1007/BF00994018.

Crenshaw, M. (2001). Counterterrorism policy and the political process. *Studies in conflict and terrorism*, *24*(5), 329-337.

Cuesta, Á., Barrero, D. F., & R-Moreno, M. D. (2014). A Framework for massive Twitter data extraction and analysis. *Malaysian Journal of Computer Science*, *27*(1), 50-67.

Dadvar, M., Trieschnigg, D., Ordelman, R., & de Jong, F. (2013, March). Improving cyberbullying detection with user context. *In European Conference on Information Retrieval (pp. 693-696). Springer, Berlin, Heidelberg.*

Dilrukshi, I., De Zoysa, K., & Caldera, A. (2013, April). Twitter news classification using SVM. *In 2013 8th International Conference on Computer Science & Education (pp. 287-291). IEEE.*

Dreiseitl, S., & Ohno-Machado, L. (2002). Logistic regression and artificial neural network classification models: a methodology review. *Journal of biomedical informatics*, *35*(5-6), 352-359.

Duwairi, R. M., Marji, R., Sha'ban, N., & Rushaidat, S. (2014, April). Sentiment analysis in Arabic Tweets. In *Information and communication systems*, 2014 5th international conference on (pp. 1-6). IEEE.

Ellis, B. (2003). Countering complexity: An analytical framework to guide counter-terrorism policy-making. *Journal of Military and Strategic Studies*, *6*(1)., 1 – 10.

Ferrara, E., Wang, W.Q., Varol, O., Flammini, A. & Galstyan, A., 2016, November. Predicting online extremism, content adopters, and interaction reciprocity. In *International conference on social informatics* (pp. 22-39). Springer, Cham.

Fisher, A., Prucha, N., Kaati, L., Omer, E., & Prucha, N. (2015). Detecting Jihadist Messages on Twitter. In *2015 European Intelligence and Security Informatics Conference Detecting* (pp. 161–164).

Fortna, V (2015). "Do Terrorists Win? Rebels' Use of Terrorism and Civil War Outcomes". International Organization. 69 (3): 519–556.

Fuchs, C. (2017). *Social media: A critical introduction*. Sage.

Johansson, F., Kaati, L. and Sahlgren, M., 2016. Detecting linguistic markers of violent extremism in online environments. In *Combating Violent Extremism and Radicalization in the Digital Era* (pp. 374-390). IGI Global.

Hosken, A. (2015). *Empire of fear: Inside the Islamic state*. Oneworld Publications.

Iskandar, B. (2017). Terrorism detection based on sentiment analysis using machine learning. *Journal of Engineering and Applied Sciences, 12(3), 691-698.*

Gates, S., & Podder, S. (2015). Social media, recruitment, allegiance and the Islamic State. *Perspectives on Terrorism, 9*(4), 107-116.

Ghajar-Khosravi, S., Kwantes, P., Derbentseva, N., & Huey, L. (2016). Quantifying salient concepts discussed in social media content: A case study using Twitter content written by radicalized youth. *Journal of Terrorism Research*, *7*(2), 10-47.

Go, A., Bhayani, R., & Huang, L. (2009). Twitter sentiment classification using distant supervision. CS224N Project Report, Stanford, Retrieved from https://cs.stanford.edu/people/alecmgo/papers/TwitterDistantSupervision09.pdf

Graves, A., & Schmidhuber, J. (2005). Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Networks, 18*(5), 602-610.

Guo, Y., Shao, Z., & Hua, N. (2010). Automatic text categorization based on content analysis with cognitive situation models. *Information Sciences*, *180*(5), 613–630.

Hong, L., & Davison, B. D. (2010, July). Empirical study of topic modeling in twitter. In *Proceedings of the first workshop on social media analytics* (pp. 80-88). ACM.

Hunter, S. (2017). Iran's Policy Toward the Persian Gulf: Dynamics of Continuity and Change. In *Security and Bilateral Issues between Iran and its Arab Neighbours*. Palgrave Macmillan, Cham.

Indra, S. T., Wikarsa, L., & Turang, R. (2016, October). Using logistic regression method to classify tweets into the selected topics. In *2016 International Conference on Advanced Computer Science and Information Systems (ICACSIS)* (pp. 385-390). IEEE.

Ikonomakis, M., Kotsiantis, S., & Tampakas, V. (2005). Text classification using machine learning techniques. *WSEAS transactions on computers, 4(8), 966-974*.

Issac, B., & Israr, N. (Eds.). (2014). *Case Studies in Secure Computing: Achievements and Trends*. CRC Press.

Jin, L., Bian, Z., Li, X., Pan, H., & Xia, S. (2009, October). Online real AdaBoost with co-training for object tracking. In *MIPPR 2009: Automatic Target Recognition and Image Analysis* (Vol. 7495, p. 749503). International Society for Optics and Photonics.Joachims, T. (1998, April). Text categorization with support vector machines: Learning with many relevant features. In *European conference on machine learning* (pp. 137-142). Springer, Berlin, Heidelberg.

Kaati, L., Omer, E., Prucha, N., & Shrestha, A. (2015, November). Detecting multipliers of jihadism on twitter. *In 2015 IEEE International Conference on Data Mining Workshop (ICDMW) (pp. 954-960). IEEE.*

Kalpakis, G., Tsikrika, T., Gialampoukidis, I., Papadopoulos, S., Vrochidis, S., & Kompatsiaris, I. (2018). Analysis of Suspended Terrorism-Related Content on Social Media. *In Community-Oriented Policing and Technological Innovations (pp. 107-118). Springer, Cham.*

Karamizadeh, S., Abdullah, S. M., Halimi, M., Shayan, J., & Rajabi, M. javad. (2014). *Advantage and drawback of support vector machine functionality. 2014 International Conference on Computer, Communications, and Control Technology (I4CT).*

Kégl, Balázs (20 December 2013). "The return of AdaBoost.MH: multi-class Hamming trees". arXiv:1312.6086 [cs.LG].

Khan, A., Asghar, M. Z., Ahmad, S., & Kundi, F. M. (2014). A review of feature extraction in sentiment analysis. *Journal of Basic and Applied Scientific Research*, *4*(3), 181-186.

Klausen, J. (2015). Tweeting the Jihad: Social media networks of Western foreign fighters in Syria and Iraq. *Studies in Conflict & Terrorism*, *38*(1), 1-22.

Kundi, F. M., Ahmad, S., Khan, A., & Asghar, M. Z. (2014). Detection and scoring of internet slangs for sentiment analysis using SentiWordNet. *Life Science Journal*, *11*(9), 66-72.

Kwon, K. H., Chadha, M., & Pellizzaro, K. (2017). Proximity and Terrorism News in Social Media: A Construal-Level Theoretical Approach to Networked Framing of Terrorism in Twitter. *Mass Communication and Society*, *20*(6), 869-894.

Lee, K., Agrawal, A., & Choudhary, A. (2013, August). Real-time disease surveillance using twitter data: demonstration on flu and cancer. *In Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 1474-1477). ACM.*

Magdy, W., Darwish, K., & Abokhodair, N. (2015). Quantifying Public Response towards Islam on Twitter after Paris Attacks. *arXiv preprint arXiv:1512.04570*.

Magdy, W., Darwish, K., & Weber, I. (2015). # FailedRevolutions: Using Twitter to study the antecedents of ISIS support. *arXiv preprint arXiv:1503.02401.*

Maron, M. E. (1961). "Automatic Indexing: An Experimental Inquiry" (PDF). Journaln of the ACM. 8 (3): 404–417. doi:10.1145/321075.321084.

Mirani, T. B., & Sasi, S. (2016, December). Sentiment analysis of ISIS related Tweets using Absolute location. *In 2016 International Conference on Computational Science and Computational Intelligence (CSCI) (pp. 1140-1145). IEEE.*

Mittal, A., & Goel, A. (2012). Stock prediction using twitter sentiment analysis. *Standford University, CS229 (2011 http://cs229. stanford. edu/proj2011 / GoelMittal-StockMarketPredictionUsingTwitterSentimentAnalysis. pdf), 15.*

Mubarak, H., Darwish, K., & Magdy, W. (2017, August). Abusive language detection on Arabic social media. *In Proceedings of the First Workshop on Abusive Language Online (pp. 52-56).*

Mohamed, E. A. S. (2016) Employment of Social Media in Response to The Terrorist Phenomena Descriptive Analytical Study on Facebook-Twitter-YouTube. *International Journal of Latest Research in Science and Technology, 5*(4), 23 – 28.

Montejo-Ráez, A., Martínez-Cámara, E., Martín-Valdivia, M. T., & Urena-Lopez, L. A. (2012, July). Random walk weighting over sentiwordnet for sentiment polarity detection on twitter. In *Proceedings of the 3rd Workshop in Computational Approaches to Subjectivity and Sentiment Analysis* (pp. 3-10). Association for Computational Linguistics.

Moon, B., McCluskey, J. D., & McCluskey, C. P. (2010). A general theory of crime and computer crime: An empirical test. *Journal of Criminal Justice*, *38*(4), 767-772.

Mujtaba, G., Shuib, L., Raj, R. G., Rajandram, R., Shaikh, K., & Al-Garadi, M. A. (2017). Automatic ICD-10 multi-class classification of cause of death from plaintext autopsy reports through expert-driven feature selection. *PloS one*, *12*(2), e0170242.

Neethu, M. S., & Rajasree, R. (2013, July). Sentiment analysis in twitter using machine learning techniques. *In 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT) (pp. 1-5). IEEE.*

Neha Upadhyay, & Prof. Angad Singh (2016). Sentiment Analysis on Twitter by using Machine Learning Technique. *International Journal for Research in Applied Science & Engineering Technology.*

Nishii, R., & Eguchi, S. (2005). Supervised image classification by contextual AdaBoost based on posteriors in neighborhoods. *IEEE Transactions on Geoscience and Remote Sensing*, *43*(11), 2547-2554.

Nock, R., & Nielsen, F. (2007). A Real generalization of discrete AdaBoost. *Artificial Intelligence*, *171*(1), 25-41.

Odeh, A. La terminologie du terrorisme d'origine arabe: Formation terminologique et rétrotraduction.O'Hara, K., & Stevens, D. (2015). Echo chambers and online radicalism: Assessing the Internet's complicity in violent extremism. *Policy & Internet, 7*(4), 401-422.

Oh, O., Agrawal, M., & Rao, H. R. (2011). Information control and terrorism: Tracking the Mumbai terrorist attack through twitter. *Information Systems Frontiers*, *13*(1), 33-43.

Omer, E. (2015). Using machine learning to identify jihadist messages on Twitter. Retrieved from https://www.diva-portal.org/smash/get/diva2:846343/FULLTEXT01.pdf

Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up ? Sentiment Classification using Machine Learning Techniques. In *Conference on Empirical Methods in Natural Language Processing, Philadelphia* (pp. 79–86).

Perdana, R. S., & Pinandito, A. (2018). Combining Likes-Retweet Analysis and Naive Bayes Classifier within Twitter for Sentiment Analysis. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, *10*(1-8), 41-46.

Posadas-Durán, J. P., Markov, I., Gómez-Adorno, H., Sidorov, G., Batyrshin, I., Gelbukh, A., & Pichardo-Lagunas, O. (2015). Syntactic n-grams as features for the author profiling task. *Working Notes Papers of the CLEF.*

Prieto-Merino, D., Macedo, A. F., Taylor, F. C., Casas, J. P., Adler, A., & Ebrahim, S. (2014). Unintended effects of statins from observational studies in the general population: systematic review and meta-analysis. *BMC Medicine*, *12*(1), 51.

Ranco, G., Aleksovski, D., Caldarelli, G., Grčar, M., & Mozetič, I. (2015). The effects of Twitter sentiment on stock price returns. *PloS one*, *10*(9), e0138441.

Rapport, M. (2015) The French Revolution and early European revolutionary terrorism. 63-76. Retrieved from eprints.gla.ac.uk/123451.

Raschka, S. (2015). *Python machine learning*. Packt Publishing Ltd.

Rauchfleisch, A., Artho, X., Metag, J., Post, S., & Schäfer, M. S. (2017). How journalists verify user-generated content during terrorist crises. Analyzing Twitter communication during the Brussels attacks. *Social Media+ Society, 3(3), 2056305117717888.*

Ravi, K., & Ravi, V. (2015). A survey on opinion mining and sentiment analysis: tasks, approaches and applications. *Knowledge-Based Systems, 89,* 14-46.

Ribeiro, P. L., Weigang, L., & Li, T. (2015, July). A unified approach for domain-specific tweet sentiment analysis. In *Information Fusion (Fusion), 2015 18th International Conference on* (pp. 846-853). IEEE.

Riloff, E., Qadir, A., Surve, P., De Silva, L., Gilbert, N., & Huang, R. (2013, October). Sarcasm as contrast between a positive sentiment and negative situation. *In Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing (pp. 704-714).*

Rish, I. (2001, August). An empirical study of the naive Bayes classifier. In *IJCAI 2001 workshop on empirical methods in artificial intelligence* (Vol. 3, No. 22, pp. 41-46). New York: IBM.

Rohit, J. (2016, August). Sentiment analysis of twitter data using machine learning technique. *In computer science and Engineering Department Thapar University*.

Rosenthal, S., Farra, N., & Nakov, P. (2017, August). SemEval-2017 task 4: Sentiment analysis in Twitter. *In Proceedings of the 11th international workshop on semantic evaluation (SemEval-2017) (pp. 502-518).*

Rowe, M., & Saif, H. (2016, May). Mining Pro-ISIS Radicalisation Signals from Social Media Users. In *ICWSM* (pp. 329-338).

Sakaki, T., Okazaki, M., & Matsuo, Y. (2010, April). Earthquake shakes Twitter users: real-time event detection by social sensors. *In Proceedings of the 19th international conference on World wide web (pp. 851-860). ACM.*

Salisbury, P. (2015). Yemen and the Saudi–Iranian 'Cold War'. Research Paper, Middle East and North Africa Programme, Chatham House, the Royal Institute of International Affairs, 11. Retrieved from https://cdn.mashreghnews.ir/ old/files/fa/ news/1393/12/10/924869_652.pdf.

Salloum, S. A., AlHamad, A. Q., Al-Emran, M., & Shaalan, K. (2018). A Survey of Arabic Text Mining. In *Intelligent Natural Language Processing: Trends and Applications* (pp. 417–431). Springer, Cham.

Salmi,Abdul Latif (2016). The violence of language in the speech of extremist organizations "ISIS" a model of research in semantic and rhetorical mechanisms. *Journal of Semiconductors (pp. 76-95).*

Selamat, A. Ã., & Ng, C. C. (2011). Arabic script web page language identifications using decision tree neural networks. *Pattern Recognition*, *44*, 133–144.

Scanlon, J. R., & Gerber, M. S. (2014). Automatic detection of cyber-recruitment by violent extremists. *Security Informatics*, *3*(1), 5.

Signorini, A., Segre, A. M., & Polgreen, P. M. (2011). The use of Twitter to track levels of disease activity and public concern in the US during the influenza A H1N1 pandemic. *PloS one*, *6*(5), e19467.

Stern, J., Berger, J. M., & Porter, R. (2015). *ISIS: The state of terror* (Vol. 7). London: William Collins.

Sun, A., Naing, M. M., Lim, E. P., & Lam, W. (2003, June). Using support vector machines for terrorism information extraction. *In International Conference on Intelligence and Security Informatics(pp. 1-12). Springer, Berlin, Heidelberg.*

Tang, J., Nobata, C., Dong, A., Chang, Y., & Liu, H. (2015, June). Propagation-based sentiment analysis for microblogging data. In *Proceedings of the 2015 SIAM International Conference on Data Mining* (pp. 577-585). Society for Industrial and Applied Mathematics.

Van Ginkel, B. T. (2015). *Responding to cyber jihad: Towards an effective counter-narrative*. International Centre for Counter-Terrorism.

Viola, P., & Jones, M. J. (2004). Robust real-time face detection. *International journal of computer vision*, 57(2), 137-154.

Von Knop, K. (2007). The female jihad: Al Qaeda's women. *Studies in Conflict & Terrorism*, *30*(5), 397-414.

Wadhwa, P., & Bhatia, M. P. S. (2014). Discovering hidden networks in online social networks. *International Journal of Intelligent Systems and Applications*, *6*(5), 44 – 54.

Wang, H., Can, D., Kazemzadeh, A., Bar, F., & Narayanan, S. (2012, July). A system for real-time twitter sentiment analysis of 2012 US presidential election cycle. In *Proceedings of the ACL 2012 System Demonstrations* (pp. 115-120). Association for Computational Linguistics.

Walker, SH; Duncan, DB (1967). "Estimation of the probability of an event as a function of several independent variables". *Biometrika. 54 (1/2): 167–178. doi:10.2307/2333860.*

Waseem, Z., & Hovy, D. (2016, June). Hateful symbols or hateful people? predictive features for hate speech detection on twitter. In *Proceedings of the NAACL student research workshop* (pp. 88-93).

Weimann, G. (2015). *Terrorism in cyberspace: The next generation*. New York City: Columbia University Press.

Torok, R. (2013). Developing an explanatory model for the process of online radicalisation and terrorism. *Security Informatics*, *2*(1), 6.

Witmer, E. W. (2016). Terror on Twitter: A Comparative Analysis of Gender and the Involvement in Pro-Jihadist Communities on Twitter. Retrieved from https://ir.lib.uwo.ca/cgi/viewcontent.cgi?article=1008&context=sociology_masrp

Wolpert, D. H., & Macready, W. G. (1997). No free lunch theorems for optimization. *IEEE transactions on evolutionary computation*, *1*(1), 67-82.

Yang, M., Kiang, M., Ku, Y., Chiu, C., & Li, Y. (2011). Social media analytics for radical opinion mining in hate group web forums. *Journal of homeland security and emergency management*, *8*(1).

Yates, D., & Paquette, S. (2010, October). Emergency knowledge management and social media technologies: A case study of the 2010 Haitian earthquake. In *Proceedings of the 73rd ASIS&T Annual Meeting on Navigating Streams in an Information Ecosystem-Volume 47* (p. 42). American Society for Information Science.

Yong, S. T., Gates, P., & Harrison, I. (2016). Digital native students–where is the evidence. *The Online Journal of New Horizons in Education*, *6*(1), 46-58.

Zanuddin, H., & Alyousef, Y. (2018). The Impact of online short and motivational videos by ISIS on Twitter towards the Saudi youth? *International Journal of Engineering & Technology*, 7(2.29), 136-139.

Zhang, X., Zhao, J., & LeCun, Y. (2015). Character-level convolutional networks for text classification. In *Advances in neural information processing systems* (pp. 649-657).