Order No:................
Serial No:................

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de La Recherche Scientifique

UNIVERSITE ECHAHID HAMMA LAKHDAR - EL OUED

FACULTÉ DES SCIENCES EXACTE

DEPARTEMENT D'INFORMATIQUE

Thèse de

Doctorat LMD

Domaine : Mathématiques et Informatique

Domaine: Informatique

Spécialité : Intelligence Artificielle

Titre

Une approche intelligente pour la détection de l'extrémisme à travers les réseaux sociaux

Par: BERHOUM Adel

- Composition du jury :

Dr. BEGGAS Mounir,	MCA,	Université d'El Oued,	Président.
Pr. MEFTAH Mohammed Charaf Eddine,	Professeur,	Université d'El Oued,	Rapporteur.
Dr. BOUZENADA Mourad,	MCA,	Université de Constantine 2,	Examinateur.
Pr. AMREOUN Mohammed	Professeur,	Université de Tebessa,	Examinateur.
Pr. LAOUAR Mohammed R´edha,	Professeur,	Université de Tebessa,	Examinateur.
Dr. BOUCHERIT Ammar,	MCA,	Université d'El Oued,	Examinateur.

El-Oued, Algérie. Février 2023

Order No: ....
Serial No: ....

People's Democratic Republic of Algeria
Ministryof Higher Education and Scientific Research

ECHAHID HAMMA LAKHDAR UNIVERSITY - EL OUED

FACULTY OF EXACT SCIENCES

COMPUTER SCIENCE DEPARTMENT

Thesis of

LMD Doctorate

Field: Mathematics and Computer Science
Domain: IT
Specialty: Artificial Intelligence

Title

An Intelligent Approach for the Detection of Extremism through Social Networks

By: BERHOUM Adel

- Board of Examiners:

Dr. BEGGAS Mounir,	MCA,	University of El Oued,	President.
Pr. MEFTAH Mohammed Charaf Eddine,	Professeur,	University of El Oued,	Rapporteur.
Dr. BOUZENADA Mourad,	MCA,	University of Constantine 2,	Examinateur.
Pr. AMREOUN Mohammed	Professeur,	University of Tebessa,	Examinateur.
Pr. LAOUAR Mohammed R´edha,	Professeur,	University of Tebessa,	Examinateur.
Dr. BOUCHERIT Ammar,	MCA,	University of El Oued,	Examinateur.

El-Oued, Algeria. February 2023

Acknowledgements

First,thanks and praise be to ***ALLAH*** *at the beginning and the end. I thank him for allowing me to overcome all difficulties. I will continue to thank you until I meet you,* ***ALLAH.***

I thank my esteemed professor, supervisor ***Dr. MEFTAH Mohammed Charaf Eddine*** *and my professors (Prof. Khaladi MKD, Dr.Lauid AK, Dr.Hammoudeh M, Dr.Mudeleh S, ... etc.) from whom we learned that success has secrets and that the impossible is achieved through teamwork and joint work, and that inspiring ideas need someone It is implanted in our minds. Their guidance and advice carried us through all the stages of writing my project. I would also like to thank the honorable chair and panelists for making my defense an interesting and special moment and for your kind suggestions and comments. Thank you.*

I would also like to extend a special thanks to my ***Wife and Family*** *for their continued support and understanding as I conduct my research and write my project. And everyone who supported me. Thank you all.*

Dedication

I dedicate this humble deed to all mankind.

To the spirit of those who encouraged me to persevere and patience throughout my life, to the most prominent man in my life

My dear father, *may God have mercy on him*

To the one with whom I sublimate the heights, and on whom I rest, to the giving heart

My beloved mother.

To my wife the highest symbols of sincerity and loyalty and companion on the trail.

To my children for the same livers.

To My brothers, my sister and family, to my friends and colleagues.

To everyone who contributed and helped me in my study journey.

Contents

List of Figures

List of Tables

Abstract

An Intelligent Approach for the Detection of Extremism through Social Networks

Extremism in its many forms is a growing threat worldwide and a threat to public safety and national security. The most prominent form is extremism in the religious text. Religious texts are among the essential parts of a people's cultural heritage and often influence societies greatly. Unfortunately, misconceptions and misconceptions can radicalize some religious fanatics and fanatics. Modern social networks allow people to express themselves, share their opinions, and show their affiliations and beliefs on many topics. Creates data in many forms, such as images, videos, and text. We used supervised machine learning (ML) to classify and analyze the textual data set of extremism from social networks at three levels for its pre-processing and text feature extraction. In this thesis, we focus on three types of classification: extremist text classification, extremist text type classification, and religious extremist text classification. During our work, we compare and discuss the results of sixteen (16) machine learning classifiers to classify a text data set for extremism after pre-processing. The first model achieved (97%) accuracy, the second model achieved (93.6%) accuracy, and the third model achieved (93.77%) accuracy.

Keywords: Social Networks, Machine Learning, Natural Language Processing, Extremism, Features, Text Mining, Religions, Classification.

Résumé

Une Approche Intelligente pour la Détection de L'extrémisme à Travers les Réseaux Sociaux

L'extrémisme sous ses nombreuses formes est une menace croissante dans le monde entier et une menace pour la sécurité publique et la sécurité nationale. La forme la plus importante est l'extrémisme dans le texte religieux. Les textes religieux font partie des éléments les plus importants du patrimoine culturel d'un peuple et influencent souvent considérablement les sociétés. Malheureusement, les idées fausses et les idées fausses peuvent radicaliser certains fanatiques et fanatiques religieux. Les réseaux sociaux modernes offrent une plateforme permettant aux gens de s'exprimer, de partager leurs opinions et de montrer leurs affiliations et leurs croyances sur de nombreux sujets. Cela crée des données sous de nombreuses formes telles que des images, des vidéos et du texte.

Nous avons utilisé l'apprentissage automatique supervisé (ML) pour classer et analyser l'ensemble de données textuelles de l'extrémisme des réseaux sociaux à trois niveaux pour son prétraitement et l'extraction des caractéristiques du texte. Dans cette thèse, nous nous concentrons sur trois types de classification : la classification des textes extrémistes, la classification des types de textes extrémistes et la classification des textes extrémistes religieux. Au cours de notre travail, nous comparons et discutons les résultats de seize (16) classificateurs d'apprentissage automatique pour classer un ensemble de données textuelles pour l'extrémisme après prétraitement. Le premier modèle a atteint une précision (97%), le deuxième modèle a atteint une précision (93,6%) et le troisième modèle a atteint une précision (93,77%).

Mots clés : Réseaux sociaux, Machine Learning, Traitement du langage naturel, Extrémisme, Classification, Text Mining, Religions, Classification.

خلاصة

نــهج ذكي للكشف عن التطرف من خلال الشبكات الاجتماعية

يمثل التطرف بأشكاله العديدة تـهديدًا متزايدًا في جميع أنحاء العالم ويمثل تـهديدًا للسلامة العامة والأمن القومي. وأبرز أشكاله هو التطرف في النص الديني. تعتبر النصوص الدينية من أهم أجزاء تراث الثقافات لدى الشعوب، وغالبًا ما تؤثر على المجتمعات بشكل كبير. لسوء الحظ، يمكن للمفاهيم الخاطئة والمظللة أن تجعل بعض المتعصبين والمتشددين الدينيين متطرفين. توفر الشبكات الاجتماعية الحديثة منصة للأشخاص للتعبير عن أنفسهم ومشاركة آرائهم وإظهار انتماءاتهم ومعتقداتـهم في العديد من الموضوعات. يؤدي هذا إلى إنشاء بيانات بأشكال عديدة مثل الصور ومقاطع الفيديو والنص.

استخدمنا التعلم الآلي الخاضع للإشراف (ML) لتصنيف وتحليل مجموعة البيانات النصية للتطرف من الشبكات الاجتماعية على ثلاثة مستويات للمعالجة المسبقة لها واستخراج ميزات النص. في هذه الأطروحة، نركز على ثلاثة أنواع من التصنيف: تصنيف النص المتطرف، تصنيف نوع النص المتطرف وتصنيف النص الديني المتطرف. خلال عملنا نقارن ونناقش نتائج ستة عشر (16) مصنفًا للتعلم الآلي لتصنيف مجموعة بيانات نصية للتطرف بعد المعالجة المسبقة. حقق النموذج الأول دقة (97٪) ، حقق النموذج الثاني دقة (93.6٪) ، وحقق النموذج الثالث دقة (93.33٪).

الكلمات المفتاحية: الكلمات المفتاحية: الشبكات الاجتماعية، التعلم الآلي، معالجة اللغة الطبيعية، النص المتطرف، الملامح، التنقيب عن النصوص، الأديان، التصنيف، التطرف.

Chapter 1

General introduction

1.1 Research scope

Social network platforms are among the largest sources of digital data in multiple forms (text, images, audio, and video), which provides a broad scope for the study and analysis of this data in many fields [2], especially the national security aspect to protect individuals and groups from abusive and extremist behavior.

Communication networks are the best place for free expression, discussion of frank opinions, and sharing beliefs in various ways (encouragement, intimidation, temptation, and intellectual recruitment) by exploiting all the social, political, and security conditions of individuals and societies. Any expression can cross borders regardless of all differences in languages, creeds, and cultures between individuals to reach an unlimited audience easily and quickly [3; 4].

Blue space facilitates the language of dialogue and discussion to express and exchange ideas and opinions. Most of its users use English to communicate with each other [5], or its applications simultaneously translate the expression.

1.2 Research problems and challenges

The dark side of social networks has led to many shocking and impactful events that have provoked and agitated many abusive phenomena (bullying, scandals, illegal human trafficking, recruitment of individuals for abusive acts, extremism, terrorism, and other abusive practices) [6; 7].

Social networks have contributed significantly to the dissemination of extremist ideas and ideologies, especially hate speech and extremist texts with fanatical and extremist religious, political, or intellectual ideas and opinions that violate all laws, and legislation, to employ individuals and groups under the banner of extremist and radical thought. Therefore, governments have had to confront such practices with various security, scientific and other agencies.

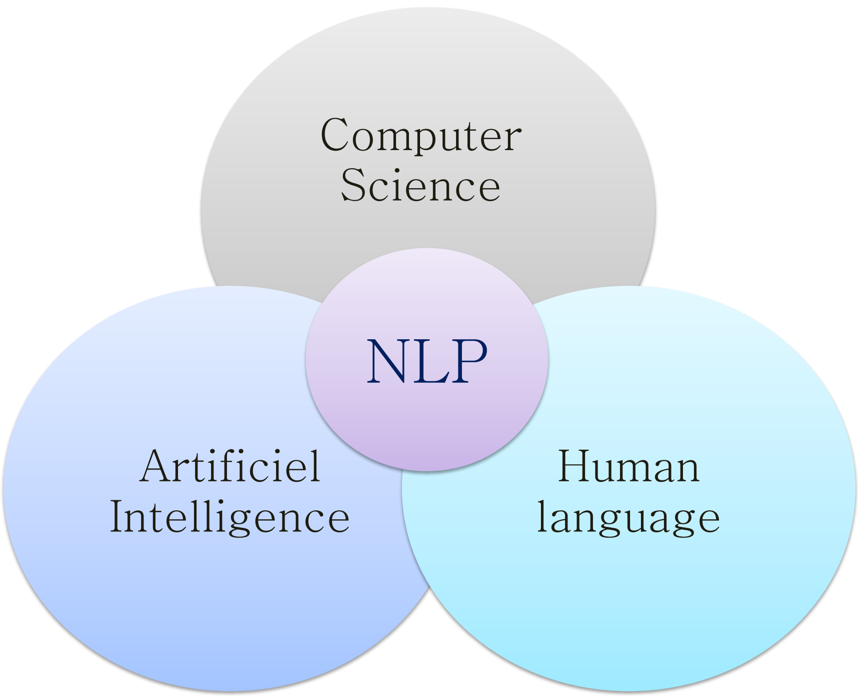

Figure 1.1: NLP is the synthesis of computer science and artificial intelligence into human language

The use of computer science and artificial intelligence contributes to natural language processing(1.1), so the scientific research of artificial intelligence tends to process text content (religious, political, intellectual, economic, and other content) through social networks to analyze the thoughts and sentiment of users to detect extremism and classify its forms.

1.3 The thesis objectives

NB : The main objective of this thesis is to find an intelligent approach that improves the results of evaluation metrics for machine learning classifiers to detect and classify extremist text types and extremist religious texts classification through social networks.

This thesis deals with the use of NLP techniques to produce a data set through pre-processed social networks to teach and test a set of machine learning classifiers to generate an intelligent model applied to similar texts to detect the intentions of the extremist text and classify it according to the different types of extremism (religious, political, intellectual and other types if found) [8].

Artificial Intelligence (AI) provides a set of machine learning algorithms to classify texts into different shapes; Through text mining (knowledge mining), a specialized form of data mining [9], a form of computer processing where knowledge is extracted through attributes (standards) in human production texts. In practice, this means modeling linguistic theories into computer systems for learning, statistics and technology to understand natural language.

Access to a pre-processed data set facilitates the work of natural language processing techniques for feature extraction, generation of new knowledge, and access to accurate and reliable results of machine learning techniques. The problem is how to direct the concept of the general appearance of the dataset using pre-processing functions (encoding and cleaning: removing strange or useless characters for text URLs... such as symbols, punctuation, hash marks, or other special characters). The thesis focuses on answering the following research questions:

Q_1: Does the Internet (Sit Web and social network sites) provide a text dataset that can be exploited directly in detecting and classifying knowledge on various topics without needing pre-processing?

Q_2: Is traditional pre-processing (with custom cleaning functions) of a dataset sufficient to process and organize it to improve the performance of NLP applications and achieve the desired results from ML and/or DL classifiers?

Q_3: Is the optimization of the data set pre-processing functions sufficient and adequate to ensure a trained and present model to give accurate results related to the core of the knowledge required from the dataset?

1.4 The thesis approach

In this thesis, the theoretician envisions a clever approach by exploiting natural language processing techniques in AI to improve the results of metrics of machine learning classifiers for the raw (dirty) dataset, taking into account other metrics (mean _square_error (MSE), confusion matrix) depending on the. And he worked on applying the approach to the data sets of the texts of religions (Islamic, Christian, Jewish, atheist, or ordinary texts) or the data set of extremism texts (religious extremism, political extremism, and intellectual extremism) proposed in this study:

- In the case of selecting the data sets of religions, the feature dictionaries for each group are extracted and tested for learning and testing of a set of machine learning classifiers to choose the optimal model. This is in case the classifications are multiple (more than three).

- In the case of selecting the extreme text dataset, the focus is specifically on optimizing the dataset as it moves to a pre-processed dataset. During this process, the dataset performs operations in encoding, cleaning, and removal functions (Cleaning, Stop-words, Lowercasing, Lemmatization, and Stemming.). Learning and testing for a set of machine learning classifiers take place at three levels.

Our work is to improve and develop a cleanup function to direct the frequency of the dataset's features toward the subject of mining in it to detect and classify extremist discourse. To achieve this, we define the padding words in the data set using a word cloud to show the most popular terms with a function that counts the number of occurrences of each word in the data set. The new function (function) cleans it up after scanning the high-frequency words in the text-feature dictionary of the pre-defined and removing the high-frequency padding words that are not in the text-feature dictionary from the dataset, thus directing the data towards the subject of the search. To obtain the results of competitive metrics, we use machine learning algorithms to categorize the extreme text.

1.5 Contributions

The research aspires to detect the presence or absence of extremism in the extremist text by reaching an intelligent classifier with acceptable and reliable metrics. Our work is to improve and develop a cleanup function to direct the frequency of the dataset's features toward the subject of mining in it to detect and classify extremist discourse. To achieve this, we define the padding words in the data set.

To achieve all this, our contribution can be limited to two main points: theoretical and practical contributions, as follows: Break down the type of extreme text by existing categories:

1.5.1 Theoretical contributions

1. An intelligent approach that detects and categorizes the extreme text content of texts. Presents methods for collecting and pre-processing text data from various social networks and/or websites to collect, filter, and prepare a data set for automated processing to extract text features and convert them into numerical vectors. Machine learning classifiers learn and test at three levels of the data set to give the best result for the proposed metrics.

2. An intelligent approach component categorizes extremist religious texts (Christian, Jewish, atheist, Islamic and plain text) where text data is collected from various social networking sites. Each religion has its dictionaries to distinguish it and identify the features of its texts. Learn and test machine learning classifiers on the dataset in (18) tests using feature dictionaries to give the best score for the proposed metrics (accuracy).

1.5.2 Practical contributions

1. Pre-processing the first dataset texts using text data processing techniques to make it more transparent and accurate in terms of focusing its features on extremes to obtain a necessary and qualified dataset by removing the high-frequency padding words.

2. Applying an intelligent approach to natural language processing with textual data processing techniques by processing qualified datasets with TF-IDF (NLP technology) to extract text features and convert them into nu-

meric vectors. Using TF (NLP technology) with feature dictionaries to output text features and convert them to numeric vectors.

3. Implementing and evaluating sixteen machine learning intelligent classifier algorithms across the refined dataset. We obtain a very high accuracy (97%) using the SVC linear classifier for extremes detection and very high accuracy (93.6%) using the SGD classifier to classify extremes.

4. Implementing and evaluating 18 tests of four machine learning intelligent classifier algorithms across (18) different data sets according to classification. We obtain a very high accuracy (93.77%) using the SVC Linear Classifier to classify extreme religious text.

Organization of the thesis

The thesis is organized into six detailed parts as follows:

Part I : Introduction general

- Chapter 1: Introduction is an arrangement for narrating the thesis in its general form, starting with the objective of the thesis, passing through the approaches and techniques to achieve this, the contributions mentioned in it, and ending with the thesis organization scheme.

Part II : State of the art

- Chapter 2: An introduction to the most important elements of the study discussing the definition and various concepts of extremism. An introduction to artificial intelligence and computer science and their relationship to natural language to generate insight into natural language processing and its different techniques and methods for language processing in social networks.

- Chapter 3: Methods for exploiting machine learning with natural language processing on a set of text data sourced from social networks to train and test machine learning classifiers to produce an acquired model at acceptable scales. Be able to generalize and deal with similar and new cases to evaluate.

Part IV: Related works

- Chapter 4: In this chapter, we survey the efforts of the various literature and related research works and their classification. We also give an overview of the different approaches related to using different techniques for handling natural language, machine learning techniques, and deep learning. Another type looks at extremism with other approaches.

Part V: General and theoretical architecture of the thesis

- Chapter 5: Presents the theoretical contributions of the thesis by giving the general and theoretical architecture of both theoretical contributions and explaining their different stages and levels. Focus on the level at which the theoretical contribution appears.

Part VI: A detailed study of the research paper

- Chapter 6: We apply two theoretical contributions to two different data sets, the first being a dataset of extremist tweets ($29,423$), and the second is a data set of different religions (3373) comments. Create an intelligent approach to detecting text extremism, profiling its type in the first dataset, and classifying extremist religious text content in the second dataset.

The last section discusses and analyzes the results and compares them with similar work.

Part VII: Conclusions general

- Chapter 7: A comprehensive conclusion of the thesis and its details of what was reached. Gathering all this data, information, methods, and techniques to obtain the results.

  A set of conclusions represents what has been concluded from edited research papers that have participated in many international and national journals and conferences.

List of publications

Journal papers:

1. *An Intelligent Approach Based on Cleaning up of Inutile Contents for Extremism Detection and Classification in Social Networks*, ACM Transactions on Asian and Low-Resource Language Information Processing, 2022 [10].
   Authors: BERHOUM Adel, MEFTAH Mohammed Charaf Eddine, LAOUID Abdelkader, HAMMOUDEH Mohammad.

2. *Machine Learning to Classify Religious Communities and Detect Extremism through Text Tweets on Social Networks*, International Journal of Organizational and Collective Intelligence (IJOCI),2022 [11].
   Authors: BERHOUM Adel, MEFTAH Mohammed Charaf Eddine, LAOUID Abdelkader, HAMMOUDEH Mohammad.

Conference papers:

1. *An intelligent approach for Educational Levels Classification in Social Networks*, 8th International Conference on Islamic Applications in Computer Science And Technology, 26-27 Dec 2020.
   Authors: BERHOUM Adel, MEFTAH Mohammed Charaf Eddine, GOURZI Sabrina.

2. *Detecting Intellectual Extremism via TF-IDF and BERT in Social Networks*, Proceedings of the 1st national Conference on Information and Communication (CICT) February 21-22, 2022, Tamanrasset, Algeria.
   Authors: BERHOUM Adel, MEFTAH Mohammed Charaf Eddine.

Communications:

1. *Machine Learning for Religious Communities Detection on Social Networks*, IPPM'20: International Pluridisciplinary PhD Meeting (IPPM'20), University of El-Oued, February, 2020.
   Authors: BERHOUM Adel, MEFTAH Mohammed Charaf Eddine.

Chapter 2

Preliminary notions and concepts

2.1 Introduction

This chapter lists the various definitions and concepts of extremism and its overview without bias or over-prescribing by clarifying its actual concept with its counterparts, such as radicalism and terrorism, describing the essential types resulting from hate speech, such as religious extremism, political extremism, and intellectual extremism. Highlighting the use of social networks to disseminate offensive content in the form of information that can be collected for processing and exploited in data mining research (text mining).

The spread of extremist discourse and its impact on individuals and people through Text (Tweets, comments) on Social Networks highlights the overlap of artificial intelligence and computer science with natural language to generate the concept of natural language processing and its different techniques and methods. It is supported by references in the literature used in classifying and detecting extremism, as well as the latest natural language processing techniques and their use in social networks to analyze the thoughts and sentiments of their users to reach the right decision in many areas.

2.2 Definitions and concepts about extremism texts

The social networks themselves are not the problem. It is how individuals and groups exploit each other rather than actual, positive, and social communication. Social networks are misused to change opinions in many way [12], such as scandals, bullying, and incitement against various types of abusive acts such as crimes, immorality, terrorism, and extremism of all kinds (Religious, Political and Intellectual). Spreading extremist discourse in its broadest sense is one of the worst uses of mass society.

2.2.1 Definitions and concepts about extremism

Extremism is a contemporary and modern word for people, especially Islamic ones [13], and even Western countries [14; 15]. In many forms and strange to current policies and organizations. However, it isn't easy to find a clear definition or context for extremism that everyone can agree on on [16; 17].

2.2.1.1 Definitions of extremism

Differences in beliefs, cultures and even policies among peoples prevented a consensus on an academic definition of extremism [18; 19]. But most of them relied on the navigator and the following features in defining extremism (political, religious, intellectual) [20; 21]. Here are some definitions of extremism.

Linguistic definition

- **Definition:** Qublan Al-Osaimi defined extremism as:*"Extremism is a departure from the line of mediation and moderation."* [22] Translator

Idiomatic definition

- **Definition-1:** J. M. Berger presents extremism as: *"Ideological perspectives, whether incompletely articulated or exhaustive in form. Individuals can be extreme in their beliefs without being violent, although extreme views often lead to involvement in violent activities."* How it shows that: *"Terrorism is a tactic and extremism is a belief system."* [23]

- **Definition-2:** K. Sharma defined extremism as *"the extent of support for the use of violence against outgroup members based on their affiliation to achieve (religious, political, or social) objectives."* [22]

- **Definition-3:** C. Bott et al. (2009) embracing the U.S. Department of Homeland Security's Definition of radicalization as *"Embracing extremist beliefs that support violence as a method to effect societal change."* [24]

One of the main difficulties associated with extremism relates to the similar contexts between it and radicalism.

2.2.1.2 Extremism and radicalism

They are synonyms or interchangeable terms to refer to the same phenomenon, leading to a misconception because they do not mean the same thing. However, there are indeed theoretical differences that make both terms conceptually different. While there is no academic consensus on the Definition of extremism and radicalism [25], their different relationship can be summarized in three main approaches:

- The two concepts are synonymous. This can be when used inconspicuously, especially in political discourse [26].

- The two concepts are different in meaning, but radicalism is a reference to the psychological process before engaging in the ideology of extremism [27].
- The two concepts are different without a necessary relationship between them [28].

2.3 Text mining in social networks

Currently, social networks are an essential link between individuals and people. Still, their dark side, especially in the fast and free dissemination of hate speech texts [29], has made it a means to produce crime, terrorism, and extremism [30; 31]. As a result, it has become an area of research and study for preventing these practices [32].

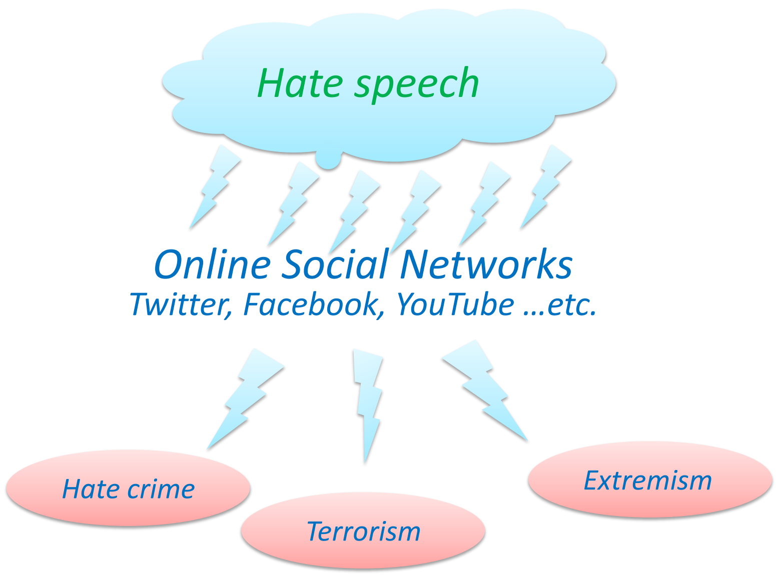

Figure 2.1: Impact of Hate speech in social networks

There are many phenomena of extremism in social networks, including pictures, videos, audio recordings, and texts. The most prevalent are texts that express the content of hatred, dissenting opinion, and dissatisfaction with the political, social, and security conditions of individuals and peoples [33; 34]. This is what saves large amounts of text data on social networks in this regard.

Text mining and text analysis allow the study of many phenomena (shopping, health, extremism, bullying, elections ... etc.) that contain a large amount of data from social networks. Facilitates the discovery and extraction of knowledge that AI algorithms used to detect and classify these phenomena in textual datasets [35]. There is a difference between text mining and text analysis:

2.3.1 Text analysis

Text analysis is one of the most widely used social science tools today to research topics in a text. Objective text analysis becomes particularly useful when it is used to predict outcomes and make inferences that are not highly accurate [36]. Many areas of use:

- Recognize irregular text data
  - Transforms unstructured data into structured data
  - Indexing and search operations
  - Natural Language Processing (NLP).
- Entity extraction
  - Classify relationships among entities
- Sentiment Analysis

2.3.2 Text mining

Text mining is a modern, interdisciplinary field that represents the intersection of information retrieval in the related fields of statistics, computational linguistics, especially data mining and machine learning [37]. It has many uses:

- Semantic determinations
- Key term (Word, sentence) identification
- Corpus of document categorization

2.4 NLP and social network analysis

The text analysis technique most used in social networks is natural language processing techniques for classification with supervised machine learning algorithms; Due to the availability of multilingual text data in social media platforms [35].

User-generated data contains vociferous content such as misspelled words, slang and abbreviations of words and sentences, and text that contains multiple languages. Moreover, comments and text tweets on social media sites are unstructured and unofficial content. The presence of low-quality content in the metadata context further complicates the problem and poses technical challenges for text mining and text analysis [38; 39].

2.4.1 NLP Techniques Used to Classify and Detect Extreme

NLP is a part of computer science that works to understand human language by using intelligence to analyze, process, and interpret it [40]. Transforms raw texts into an organized data set by capturing grammatical, semantic (features), and linguistic information to generate and infer new knowledge. In practice. In two main stages:

2.4.1.1 Data pre-processing

Data pre-processing is essential to natural language processing, as it helps to identify and generate clean and pure data for evaluation during analysis [41; 42]. This process includes a set of techniques (software functions) that allow NLP techniques to process and analyze data to convert it from NLP data into machine-understandable code. Data pre-processing techniques are:

- o **Tokenization:** It is the first process in which the sentence is divided into smaller units (called tokens).

- o **Cleanup:** Remove strange or useless characters and words such as symbols, punctuation, hashtags, and text URLs if necessary.

- o **Stop words:** Are removed because they do not carry relevant information in most contexts.

- o **Lowercase:** They are removed because NLP techniques do not distinguish between Lowercase and uppercase letters.

- o **Lemma:** Return verbs to their roots.

Transform the raw text dataset into a pre-processing and structured dataset suitable for various machine learning algorithms or analysis methods.

2.4.1.2 Feature extraction

The second essential stage of natural language processing and after pre-processing the text dataset is to use various approaches to extract text features 2.2 and transform them into symbols by capturing their lexical, syntactic, and semantic information [43]. In the end, these tokens are used as inputs to various algorithms to analyze knowledge through the given set or to gain new knowledge from it [44].

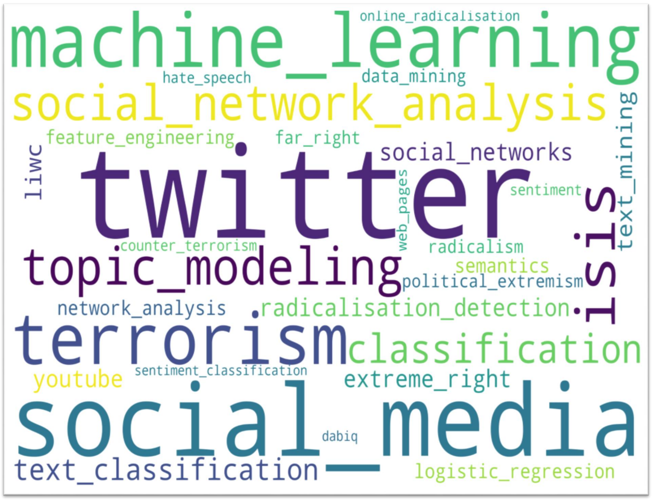

Figure 2.2: Word cloud of features extracted

There are three (3) basic approaches for extracting text features according to the type of linguistic information obtained from the pre-processing of the text data set:

- **Lexical or Vectorial Based Features:** After the tokenization process, it defines a weighting technique to compute the token (terms) appearance's frequency in a text. Exist different techniques to generate this vector representation (N-grams [45], Dictionaries [46], TF, TF-IDF [47], Dichotomous [48], Log-likelihood [49]).

- **Word embedding for Neural Language Models:** Since 2017, transformers [50] has become a dominant modern technology for achieving cutting-edge results using modern methods and techniques in various NLP tasks [51]. Practically it is the use of neural networks to transform the tokens obtained from pre-processing into meaningful vectors, to capture the relationship between them [52], and to extract information about linguistically related words, among the most prominent techniques used in classifying extremism (Word2vect [53], FastText [54], GloVe [55]).

- **Syntactic and Semantic Features:** It relies on data analysis according to context for generating features representing the text [56].

The most appropriate choice of NLP techniques in extremism research depends on the characterization of extremism from a descriptive perspective. To choose four different descriptive topics:

- Terms: descriptive analysis of the terms commonly used by extremists. Characterization of the type of extremist vocabulary.
- Topics: detection of the most common topics discussed by extremist texts.
- Sentiment: analysis of the sentiment and tone of an extremist discourse.
- Semantic: analysis of the contextual information around terms inside an extremist text.

2.5 Conclusion

This chapter explains that we find the breach of traditions and beliefs, especially religions. The breach of policies is the reason for the impossibility of a unified definition of extremism between different peoples and governments and the differences between the concepts of extremism and radicalism. How has the development of social networks had a direct, effective, and immediate impact on the dissemination of extremist content? Especially in the doubt of texts, in the considerable amount that formed textual data sets with different features and beliefs and is ready for analysis and study.

Text mining in social networks is becoming increasingly crucial for identifying irregular text data using natural language processing techniques and machine learning classifiers.

In this chapter, we find ways to use natural language processing techniques in the field of artificial intelligence and how their uses have increased in importance with the development of social networks to provide a considerable amount of data of all kinds. To reach the exact model for the detection and classification of the extreme and extreme text, the motives for adopting the correct choice of natural language processing technology with an appropriate classification algorithm on accurate and targeted data to reach the best results for the model test measures.

Chapter 3

Machine Learning for Natural Language Processing

3.1 Introduction

In the fifties of the twenty-first century, studies on natural language processing shone, and the publication of the article "Computing Machines and Intelligence" [57] by Alan Turing was the beginning of such studies. In general, natural language processing is defined as the automatic processing of a machine by natural language programs.

Before the 1980s, natural language processing techniques depended on complex manual rules and software. By the late 1980s, merging machine learning algorithms with enormous computational power with natural language processing techniques played a role in language processing. Because of it, the traditional techniques (e.g., transformational grammar by Chomskyan theories of linguistics [58]) disappeared.

This chapter presents the relationship between machine learning and natural language processing to produce a learned model. A model is a mathematical representation trained on a set of textual data to generalize and deal with similar and new cases to evaluate them [59]. The set of irregular text data extracted from social networks requires a particular approach to machine learning. This is because text data generally contains a considerable amount of features and dimensions for words and phrases [60].

Machine learning for natural and text language processing includes a set of statistical methods for identifying and extracting speech features, entities, feelings, and other associations of text words [61]. They are the model used on data sets to classify them or extract knowledge.

3.2 Machine Learning for NLP

Natural Language Processing is a machine learning application that analyzes text with intelligent algorithms(3.1) to extract knowledge from the text [62]. Text from different sources (social networks, responses to various questionnaires, financial, legal, medical documents ... etc.).

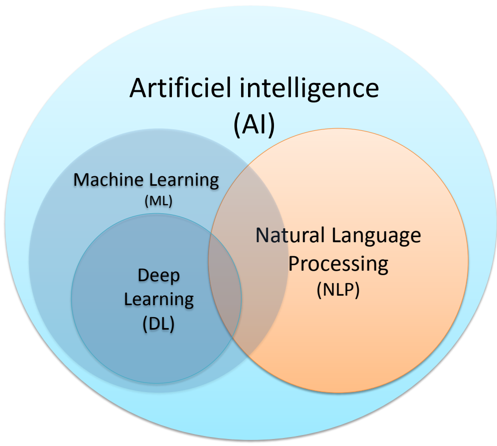

Figure 3.1: NLP in Machine Learning and AI

3.2.1 Types of Machine Learning for NLP

Machine learning is a subset of artificial intelligence whose work is predictive analytics or predictive modeling. It is "the ability of a computer to learn without explicitly programming it," Arthur Samuel termed "machine learning" in 1959. In practice, machine learning is the application of algorithms to receive and transmit an input data set to predict the output value with acceptable metric values. There are three machine learning algorithms: supervised, unsupervised, and Hybrid.

3.2.1.1 Supervised Machine Learning

Supervised Machine Learning is a statistical method for identifying categorized data sets to make predictions of text or speech [63] and sentiment to create an acquired model that is applied to other texts. During which texts are highlighted to indicate what the machine is looking for.

3.2.1.2 Unsupervised Machine Learning

Unsupervised Machine Learning is training a model without coding the text. It is often used in Clustering to group similar documents together into groups [64]. To sort these groups based on relevance and need. It has other uses (e.g., Latent Semantic Indexing (LSI)).

3.2.1.3 Hybrid Machine Learning Systems

Hybrid Machine Learning Systems is important to understand the importance and rationale for choosing supervised learning or unsupervised learning when you can get the best of both to create one effective system [65].

3.2.2 Machine learning algorithms

Knowing all the terms mentioned above makes us realize that we are dealing with machine-learning algorithms geared toward data scientists.

In addition, the machine learning algorithms used in our research are supervised machines because they train a healthy "labeled" training of text data, based on which the machine predicts pre-sorted output, generally a process that distinguishes some input data into the correct output [66]. We highlight four machine learning algorithms that are very different in dealing with data [67], and they are the most used in recent years to provide accurate results:

3.2.2.1 Support vector machines (SVMs)

A supervised algorithm that specializes in classification and regression. The data is plotted in space as a plot representing the number of features [68]. SVMs represent the coordinates of individual observations over space. Used in applications such as classification (data, facial expressions, text, steganography, speech recognition, etc.).

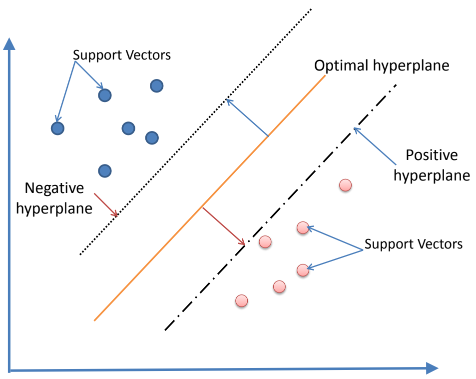

Figure 3.2: Support vector machines

The graph shows that the algorithm selects the extreme points on the vectors that help create the data's hyperplane. In extreme cases, they are called support vectors, the source of the algorithm's naming of the support vector machine. In the chart below, two different categories are categorized using decision boundary and hyperplane.

3.2.2.2 Naive Bayes algorithm

The naive Bayes algorithm is a probabilistic Machine Learning algorithm based on a Bayesian probability model, Intended for classification processing [69; 70]. Its basic premise is that changing the attributes under study does not affect the other value because they are independent.

Mathematical representation of the algorithm.
If X, Y = probability events, P(X) = probability of X being true, P(X—Y) = conditional probability of true X being true.
Bayes theorem equation:

$$P(\theta|\mathbf{D}) = P(\theta)\frac{P(\mathbf{D}|\theta)}{P(\mathbf{D})} \tag{3.1}$$

This algorithm can handle massive datasets and helps make real-time predictions [71]. Its applications include prediction, text classification, sentiment analysis, and more.

3.2.2.3 Decision trees algorithm

It is a map of the possible outcomes of a set of decisions. You are likely to expect the best option based on a mathematical structure, and it is helpful while brainstorming about a particular decision [72].

A decision tree is formed with a root node (decision node) through which the child nodes represent the possible outcomes and thus can solve other possibilities

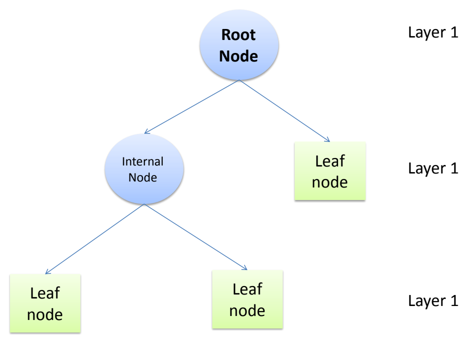

Figure 3.3: Node Decision trees algorithm

used in classification problems [73].

3.2.2.4 Random Forest algorithm

The Random Forest algorithm is a supervised algorithm composed of several different decision trees with different order samples for each decision tree.

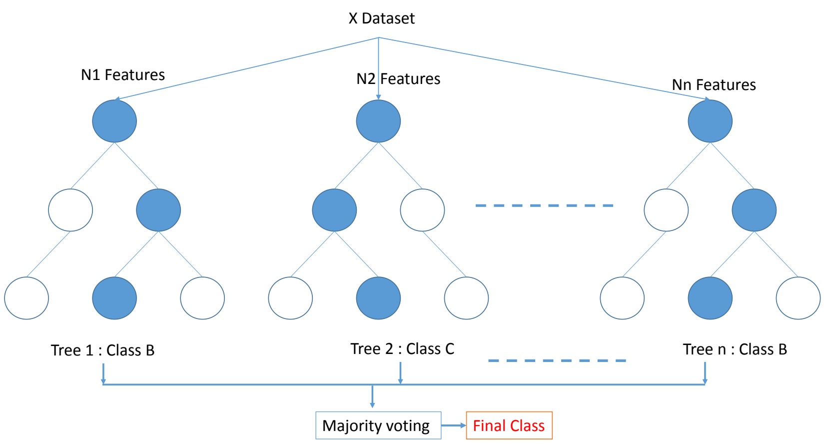

Figure 3.4: Random Forest

It maintains the integrity of the results for the loss of evidence, and its choices in the training data are random [74]. Each dataset has a decision tree and its results. The result of the prediction is the most voted on.

3.3 Conclusion

It is clear from this chapter that the intersection of artificial intelligence with natural language generates an intelligent approach capable of exploiting textual data sets that originate from social networks. An approach that gives a qualified learning and testing model to give accurate validation metrics results. One of the essential techniques of artificial intelligence is that natural language processing techniques can overlap with supervised machine learning classifiers by applying them to a set of textual data that can produce a model that is applied to similar texts to give results of accurate classification or reveal the feelings and intentions in its content.

unsupervised, and hybrid) and their "semantics." Since our research is for content detection and classification, the presentation of the four most commonly used supervised machine learning classifiers gives an idea of how to enter and output data to classifiers.

In the following, we present and detail the general structure hypothesis of this intelligent approach to give an available picture of the basic levels of the approach structure and how to apply them.

Chapter 4

Related works

4.1 Introduction

This chapter reviews the relevant literature from three axes (Figure 4.1). First: We review the common literature on the same topic of research, which is the detection of extremism in its various forms (religious, political, intellectual) or even one type of them, and we cite its proposed technical methods for processing the data set and extracting its features and the results of different measures resulting from the use of machine learning techniques and deep. Second: We review the literature that relied on the same dataset processing techniques or other network techniques in solving different problems (e.g., religious or non-religious texts). Third: a literature review that follows the same or similar approach but differs in the dataset or processing techniques to solve forms other than extremism (e.g., shopping, sports, elections).

Before starting this review, we highlight the comprehensive review (pre-2015 literature review) done by Agarwal and Sureka (2015) [75], which presented a

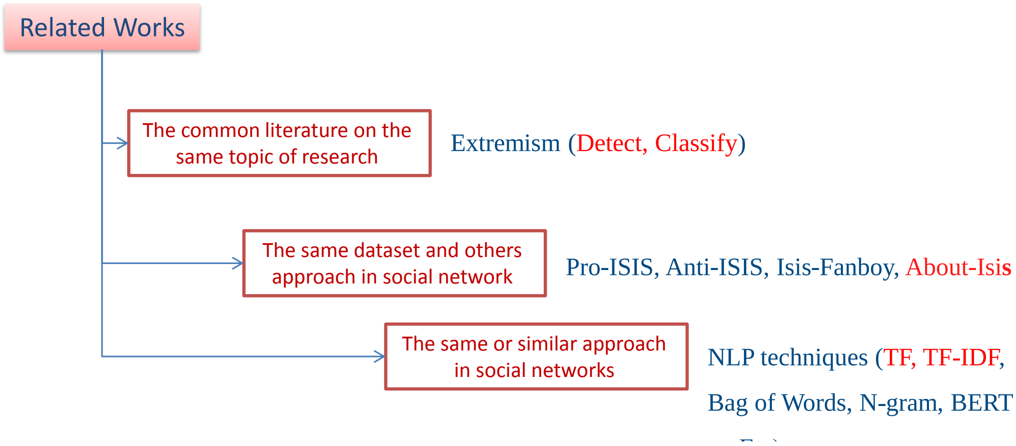

Figure 4.1: Axes of analysis and study of related literature

comprehensive and detailed survey of the literature (more than 40 scholarly articles) over ten years (2006-2015) aimed at understanding the latest developments in online social network intelligence technology to counter and combat ISIS. This work addresses the problem of automatic identification of extremism on the Internet in general and on social networks in particular. It highlights the prediction of events related to civil unrest. It was to manipulate the data by searching for keywords and highlights in documents, articles, and texts published across the following three platforms (online social media platforms, intelligence and security, and text mining). In conclusion, it was found that written texts are the most studied and analyzed data. Their results were acceptable, using various text-mining techniques, machine learning techniques, and deep learning.

4.2 The standard literature on the same topic of research

The following literature reviews the methods and methods for detecting extremism in various forms. Using natural language processing techniques and machine and deep learning classifiers.

Liu Shushu and others [76] show the importance of classifying the rhetorical article as religious or not, the impact of religious discourse on social, political, security, and other fields, and the extent of its impact on society and the individual. It also shows the difficulty of classifying religious words because they have no precise reference.

Xia Xie and others [9] applied machine learning techniques to sentiment analysis of extremist texts in social media to detect extremist user interaction. A database was created with 20,000 tweets from multiple accounts, and the system focused on three tasks: detecting extremist users, identifying texts with extremist content, and predicting. The effectiveness of other users of extremist content and the model's accuracy concerned the detection of extremists. (Taira Teemu, 2022) [77] has the same goals but differently. It implemented the Naïve Bayes algorithm with the classic feature set, the system categorizing user reviews into positive and negative categories in languages other than English.

Emilio Ferrara and others [78] present a new framework for analyzing terrorism-related content with a focus on classifying tweets as extremist and non-extremist and developed a Tweet scoring system using deep learning. This work introduced an emotional extremist ranking system based on user tweets. The experimental results showed that the proposed system achieves excellent performance compared to other methods (precision, recall, and measurement). However, the system has certain limitations.

In another study, Azizan Sofea Azrina and others [79] compared and contrasted different religious texts using point-of-sale labels and a document terminology matrix. In this work, ML was used to analyze religious texts lexically. They used common names to analyze religious texts from similar geographic locations. Religions from the same geographical area have the most names in common. This work indicates that the location influences religious texts. It also shows that some religious texts contain references to the deity of other religions, indicating that the texts of these religions compare to other religions in the same field.

The work of (Sarker Aditi et al., 2020) [80] introduces an improved system model for analyzing Twitter data and detecting terrorist attack events. K-Nearest Neighbor (KNN) and Support Vector Machine (SVM) are used to predict whether a terrorist attack has occurred or not. The author-Corasick algorithm is applied to perform pattern matching and assign weight using ternary search to find the weights of the predefined keywords. The weights are categorized into three categories: Terrorist Attack, Severe Terrorist Attack, and Normal Data, and the weights are used as attributes for ranking.

4.3 The same dataset and others approach in social network

Other literature deals with research on extremism through social networks but without the need for applications and methods of natural language processing, including:

The work (Chang Victor et al., 2022) applies social network analysis in two experiments. In the first experiment, social network analysis is conducted on students' friendship networks to find relational patterns. Then, three community detection methods are used to divide the student network. The RSiena package is used to illustrate the evolution of friendship networks. The second experiment aims to improve security issues within the Bitcoin Trust Network on the BTC-Alpha platform by applying social network analytics.

Using the homicide dataset provided by White and Rosenfeld and using Rstudio to structure the network of suspects and the network of victims and work (Chang et al., 2021) [81] aims to explore the use of Social network analysis to identify the most active suspects and possible network gang crimes. This paper concludes that the criminal gang and victim group in homicide cases could be studied by performing centrality analysis and detecting cliques in these two single-mode networks.

The work of (Jamil M Luqman et al., 2022) [82] used methods based on clustering, the Feature-pivot approach, and the document-pivot approach to detect dangerous events on social networks. The subcategories consist of dangerous events based on scenarios, feelings, and actions. Although in some cases, important events also come from social media. The usefulness of social media offers a significant advantage in initially detecting such events.

4.4 Similar approaches in social networks

The following literature addresses several disadvantages of networking using TF-IDF technology from a suite of natural language processing techniques. It uses machine learning classifiers to create a learner model capable of detecting and classifying their problems. The following literature discusses ways to use TF-IDF as a natural language processing technique. Its work is to extract text features and transform them into digital vectors. It learns and tests machine learning classifiers on digital vectors for text features to create a learner model to discover and classify their content.

Where Bhattacharjee Uddipta and others [83] have raised the problem of spreading rumors and their impact on the correct information of the actual event through social networks. In the second paper, (Islam Md Manowarul and others [84]) raised the problem of bullying, especially for women and children on social networks. They worked on designing and developing an effective technology to detect offensive messages and cyberbullying by integrating natural language processing and machine learning. Two distinct features, Bag-of-Words (BoW) and (TF-IDF) are used to train and test the four Machine Learning workbooks. They have proposed a new approach to extract features by re-measuring the TF-IDF score for some particular terms taking into account the private label information with training data on an SVM classifier. As for the last paper (Han Wenlin and others [85]), issues related to fake news that lead to misleading destinations, this work aims to compare and evaluate the various methods used to mitigate this problem and some machine learning approaches.

Alvaro de Pablo et al. [86], proposed several metrics to describe both radical and non-radical texts. Considering the results obtained in evaluating the Random Forest classifier, the mentioned style features perform poorly when using (Style or Bigrams) to extract text features with approach (TF) compared to the more detailed representation of text when combined (Bigrams + Style).

In their work, Logan Macnair and Richard Frank [87] used sentiment analysis and a natural language processing (TF) technique to show the consistency of language used in ISIS journals to highlight the perception of magazine content to ISIS enemies through extremist narratives. The results were evaluated by calculating the standard sentiment score and regression scale for all keywords (features).

Ofra Klein and Jasper Muis [88] worked on a study comparing the rhetoric of the Facebook pages of far-right parties, movements, and communities in Western Europe as a result of their limited political opportunities. Sentiment analysis was used for the most frequent words extracted by natural language processing (NLP) technology and inferred the likeness relationship between Facebook groups for far-right movements and organizations.

Ryan Scrivens et al. [89] used a sentiment-based algorithm adapted to professional criminal procedures to perform sentiment-based identification and sentiment-based temporal identification of the radical authors. The result is an assessment of how users' hostile behaviors for certain castes (Semitic, Black, White Sovereignty Forum) evolve by applying an intelligent approach to uncovering radical online dissemination pathways.

Leevia Dillon et al. [90] study focused on sentiment-based quantitative analysis of social media posts by ISIS supporters and foreign fighters during the period (2009-2015). They used a multi-method design to define qualitative content analyzes of participants across five themes (for group threats, societal grievances, the pursuit of meaningful issues, religion, and commitment). This work showed that supporters (10 out of 18; 56%) posted more radical content than fighters (5 out of 14; 36%) on social media.

4.5 Comparisons and discussions

4.5.1 Comparisons

Table (4.1) shows a literature review regarding detecting extremism in a dataset sourced from social networks. From the table, it is possible to compare the natural language processing techniques and machine learning and deep learning classifiers used on the same dataset.

It is noted that natural language processing techniques vary between different works of literature, and this is due to the need for the approach used in the feature extraction method. Still, in the end, it aims to extract the features of texts and digitize them in vectors to learn and test machine learning classifiers.

Abd-Elaal et al., [91], a proposal for an intelligent system that detects Pro-ISIS accounts by analyzing language and behavioral attributes with the formation

Table 4.1: Similar works to the proposed solution.

Work	Source dataset	Technique(s) NLP	Technique(s) ML or DL
(Abd-Elaal et al, 2020) [91]	Pro-ISIS and Anti-ISIS	TF-IDF and Skip-gram "Mazajak" [92]	Bernoulli NB, DTC, K Neighbors C, Linear SVC, LR, RFC
(Ashcroft et al., 2015) [93]	TW-PRO, TW-RAND, TW-CON	Stylometric features (S) Time based features (TBF) Sentiment based features (SF)	Classification based on features using SVM, Naive-Bayes and Adaboost
(Saif et al., 2017) [94]	Pro-ISIS and Anti-ISIS	Unigrams Features, Sentiment Features, Topic Features, Network Features	SVM classifiers
(Benigni et al, 2017) [95]	Tweets	Iterative Vertex Clustering and Classification (IVCC)	Common algorithms include K-means, Newman and Louvain Grouping
(Sharif Waqas et al., 2019) [96]	Tweets	N-gram , TF-IDF	SVM
(Mussiraliyeva et al, 2020) [97]	Tweets	Word2Vec and TF-IDF	Gradient boosting and Random forest

of two subsystems. The first is a subsystem that crawls Anti-ISIS accounts to detect ISIS accounts through social networks, and the second is a subsystem that runs on queries to discover Pro-ISIS accounts. They used (TF-IDF and Skip-gram "Mazajak") techniques to extract the features of text content with a set of machine learning classifiers. The SVC linear classifier gave acceptable results $(0,62\%,0,94\%)$.

Ashcroft et al. [93] used a machine learning approach that categorizes a tweet as containing material that supports jihadist groups. Their results were preliminary and needed more tests. The results were obtained using the limited dataset

(TW-PRO, TW-RAND, and TW-CON) using three methods of text feature extraction (Stylometric features (S) Time based features (TBF) Sentiment based features (SF)) with machine learning classifiers (SVM, Naive Bayes, and Adaboost) to advance the work of detecting extremist content on social media, in particular on Twitter. It should be noted that the approach was intuitive for detecting extremist content.

Unlike other research that relies on the lexical representation of published content in its study, Saif et al. [94] used an intelligent approach that extracts and uses the basic semantics or specific features of the words posted and used by social media users to determine their pro/anti-ISIS positions. The result is two classifiers trained on semantic features that have demonstrated superiority over models trained on lexicon, sentiment, subject, and network features. The value of the improvement in the mean scale (F1-score = 7.8%)

Benigni et al. [95] presented an iterative study of header clustering and classification (IVCC) using popular K-mean, Newman, and Louvain GroupClassification to develop a scalable analytic approach to discovering the OEC (a group of more than 22,000 Twitter users whose online behavior is in direct support of ISIS or a contributor to the group). Or they spread propaganda through retweets) in social networks.

Sharif Waqas et al. [96] analyzed the sense of cyber content of extremist text from Twitter sites because its content is short, noisy, context-dependent, and dynamic. Exploratory dataset analysis (EDA) was performed using Principal Component Analysis (PCA), with reduced features based on PCA and with a complete feature set (TF-IDF features extracted from n-gram terms in tweets). Training and testing machine learning classifiers (naïve Bayes (NB), K Nearest Neighbors (KNN), Random Forest (RF), and Support Vector Machine (SVM)) to get an average accuracy of 84

Exploring the possibilities of automatic identification of extremist content using machine learning classifiers (Gradient boosting and Random forest) after extracting text features using natural language processing techniques (Word2Vec and TF-IDF) on text content from the Internet. During this literature, (Mussiraliyeva et al.) [97] worked to uncover content that indicates a tendency to extremism in preventing extremist and terrorist activities. The acceptable value obtained for the model's accuracy is using (Gradient boosting with word2vec), which was (89%).

4.5.2 Discussions

Out of analyzing previous literary methods and their results, it is clear that the concept of extremism is different among researchers because its concept determines the features on which natural language processing techniques depend, the techniques of which differ from one literary to another. Machine learning and deep learning techniques vary, but most of the literature showed that the best machine learning algorithms that give good results for measurements in training, testing, and evaluation are (SVM, RF, DT, and NB). Deep learning (CNN, LSTM)

The literature related to the thesis topic is numerous because the research topic is complex and has many aspects. On the other hand, its most essential elements are constantly evolving and progressing (the development of social networking applications on the Web, ways of dealing with and exploiting individuals and groups with the blue world, and the development of natural language processing technologies). The analyzed literature shows and clarifies many aspects of the research on the subject. Its results were acceptable to a large extent, but they can be improved if the improved dataset is. Therefore, it has become necessary to focus on the accuracy and credibility of the results and the quality of the trained, intelligent models rather than the goal being to reach the best results, which could be better.

Therefore, the research orientation of our thesis was based on the essential element related to the accuracy and reliability of the results, namely, the thClassificationand formation of the raw data set and the pre-processing of the structured data set. Each depends on the approved concept of the features of the extremist text, based on the concept of general extremism, without any specification or orientation to determine the objective of the research.

4.6 Conclusion

In this chapter, it is clear from the relevant literature and reviews that the idea of exposing extremism is a topic that has prompted many researchers and programmers to stand up to the spread of extremism. The abusive use of social networks and their surprisingly rapid cross-border advantages contributed to its spread, where text content was the most prevalent, which necessitated the use of natural language processing techniques [98].

This chapter shows the approaches and contributions of businesses that worked on the same data set, the literature that used the same approaches and techniques in dealing with texts in social networks, and the literature that differed in natural language processing techniques and machine learning and deep learning workbooks. Other works deal with extremism through social networks without processing aggregate data but instead work on securing networks.

Therefore, research in parallel with optimization and help to achieve the best results with current and ongoing research is the right idea.

Chapter 5

The proposed approach and methodology

5.1 Introduction

This chapter discusses the hypotheses and techniques incorporated throughout the thesis's general architecture levels, from obtaining data from social networks or websites to creating a model that helps professionals and those interested in making the right decision. These statements are often unstructured and informal, and their content is noisy and obscene [38; 39]. This data passes through three levels, as shown in the general structure Figure(5.1).

We analyze the levels through which the approach is implemented. List all the tools and techniques used during each level to arrive at the proposed approach's overall shape; clarify and discuss the procedures used (techniques, algorithms, Features, dictionaries ... Etc.) step by step, we refer to some examples and literature.

The general architecture figure Figure(5.1) shows that the first and second approaches do not pass the automatic stage (pre-processing) during the first level

(L1 (6.5))of their implementation. That the third approach passes directly to the second level (L2 (6.5)) to apply the TF technique of NLP to it after extracting dictionaries of the features of religious texts, detailing it as follows:

5.2 The proposed approach

As illustrated in Figure(5.1), the general architecture of the approach shows that raw data imported from social network users pass through three levels:

- Pre-processing of raw data in several stages.
- Extracting features by choosing the appropriate data pre-processing technique.
- Suggest the most suitable machine learning classifiers based on the proposed metrics.

At each level, a set of artificial intelligence techniques are implemented, and the sequence of levels is very important. It is not possible to start a level without completely completing the level before it so that the work is integrated.

The result is two classifiers:

- The first classifier detects the existence of the action in the text content (extremist or No extremist - the text content is religious or not religious) through the dataset textual.
- The second classifies the data resulting from the first classifier into a set of class according to the required types (Type 1, Type 2. .... Type n), (Example: Types of religious text: Muslim, Christian, Jew, atheist and Buddhist. ... etc.). Below is a detailed explanation of the three proposals.

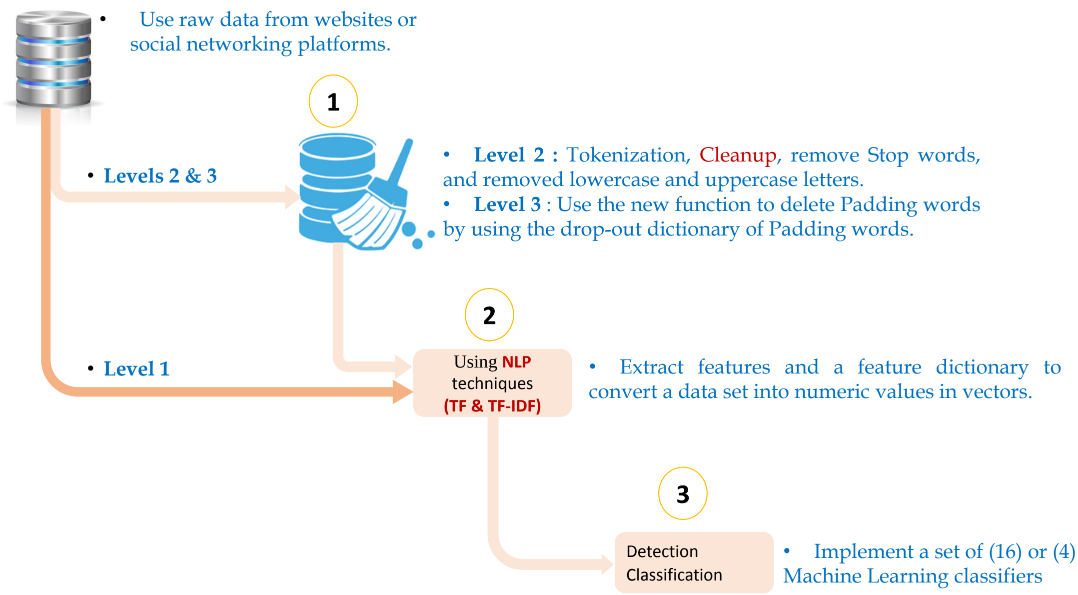

Figure 5.1: General architecture of the proposed approach

5.2.1 Dataset used

The proposed approach is based on textual data sets from social networks, specifically Twitter and Facebook, because it is the most popular [99]. Their textual content is rich in phrases and words with influential properties. The Twitter and Facebook apps allow researchers to download tweets and comments in unlimited quantities [100; 101]. Also, the collection and classification of textual data is a rather tricky task [102] because it requires the accuracy and integrity of data collection [103]. Coding and gathering data takes a lot of time and knowledge to reach the required accuracy that defines the class and genre of each tweet.

5.2.1.1 Dataset sources

There are two basic methods for acquiring raw data; the first: is directly from social network platforms through applications that collect data directly (Facebook Export comments, socially) [101; 104] This is for Facebook, and there are many

of them. For Twitter, the data source provides an additional data source that is freely available to researchers around the world to conduct research projects (The Twitter developer portal) [100; 105]. The data is in a text, Excel, or comma-separated values (CSV) file. It is easy to process [106; 107].
Second source: Websites for collecting, exploring, and coordinating the sharing of high-quality data for better and more efficient use by researchers in the data science community. Of the essential dataset finders:

- **Google Dataset Search:** It works in the same way as Google Scholar, providing all its information with a data set [108].

- The Big Bad NLP Database: Is a platform that provides datasets for natural language processing tasks [109].

- **Kaggle Datasets:** It is an open online platform for collecting datasets to publish and make it easier for data scientists and machine learning programmers to collaborate and share data and compete to solve contemporary challenges of data science [110].

Many reliable sites provide data sets, including general and specialized, making it easier for researchers and programmers to collect and organize. Nevertheless, extracting, collecting, and organizing them directly from social networking applications takes excellent time and effort [111; 112].

The dataset should categorize and order data and characterize them by different words and phrases (features). The data must also have an identifying name or the link from which it was taken and any information that is useful when referring to it.

5.2.1.2 Data pre-processing

A set of functions converts a raw dataset into a ready-made data set for natural language processing techniques. Data pre-processing phases:

Tokenization data: It can be seen in Table (5.1) that the encryption function splits the tweet into separate words, breakpoints are removed, and all letters are lowercase.

They are words or letters that do not carry relevant information and often do not constitute features of the word that are omitted, and numbers and links are deleted as in the third column of the Table(5.1).

Table 5.1: Tokenization and clean the dataset.

Tweet raw	Tweet tokenization	Clean tweet
IS but one is important to live when the other one is not!	[”, ’is’, ’but’, ’one’, ’is’, ’important’, ’to’, ’live’, ’when’, ’the’, ’other’, ’one’, ’is’, ’not’, ”]	IS important live

Removed padding words: They are words that are not related to the subject of the study (incorrect, irrelevant, or incomplete words) and have a high frequency, which appears prominent after being shown on the word cloud (Figure 5.2) in the first level of processing the raw dataset. Moreover, by applying the pre-defined cleaning functions to the raw data set, the result of the cleaning appears in the word cloud of the second level (Figure 5.3). In addition, at the third level, we create a new function that scans the dataset with a dictionary of discarded words and removes them so that the result appears in the word cloud (Figure 5.4). It is extracted from the dataset with the function of extracting all the words of the dataset in their frequency order and manually extracting the discarded words with higher frequency and then arranging them according to their frequency to create a dictionary which we named ((pad_words (6.12)).

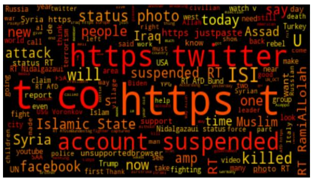

Figure 5.2: Raw dataset.

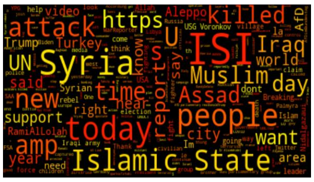

Figure 5.3: Dataset cleaning.

Figure 5.4: Dataset cleaning without padding words.

After implementing the new function of removing the padding words, the results appear in the word cloud (Figure 5.4), where words Subject-related words stand out, indicating that they are in the foreground in terms of their frequency value most common in the dataset. Hence, the dataset is purer and ready for the following stages.

5.2.2 The proposed technique NLP

Principal features include ideas that can positively or negatively influence opinions. It specifies the intention of its user to direct individuals and groups to a particular idea. It can be distinguished personalities, books that people relate to, such as religious books or charters and laws, and they can be articles by a trusted writer and inferring. Their comments and tweets on social networks make them a force to influence the sentiment of individuals and peoples [113].

Natural language processing applications work to extract these features and calculate the relationship between them (subjectivity, polarity) and their impact on the core of the topic, to infer the sentiment of each comment or tweet [114].

5.2.2.1 Sentiment analysis

Opinion mining is the analysis of text to discover and categorize personal ideas and opinions with applications of natural language processing and computational linguistics. The goal is to identify the opinion, intent, or emotional state that the author wishes to elicit in the reader [115; 116]. Objective tweets that are straightforward, such as asking, tend to be emotionless, and tweets that express character or status tend to have positive or negative feelings. An emotional state or state is determined by categorizing the subjectivity of tweets and then categorizing them into positive, normal, and negative tweets. Figure(5.5). A principle for detecting the subjectivity and polarity of texts (comments, tweets).

Subjectivity analysis: *subjectivity* is a value floating on the closed interval [0,1]. The value (0) represents indisputable factual information. The value (1) for subjectivity represents the writer's personal opinion, sympathy, or self-judgment.

Polarity analysis: *polarity* is a value floating on the closed interval [-1,1] where (1) is all-positive, (-1) is all-negative, and (0) is neutral (the number of negative features equals the number of positive features). Polarity values for a statement can range (from 1 to -1) depending on its content of distinct words 5.5.

We rely on the TextBlob package in Python as a convenient way to do many NLP tasks, mainly calculating polarity and subjectivity from which to infer tweet sentiment.

5.2.2.2 Feature extraction

Natural language processing provides several techniques for feature extraction, including Term Frequency - Inverse Document Frequency (TF-IDF) [117], whose goal is to capture the relative importance of words across texts in a corpus without

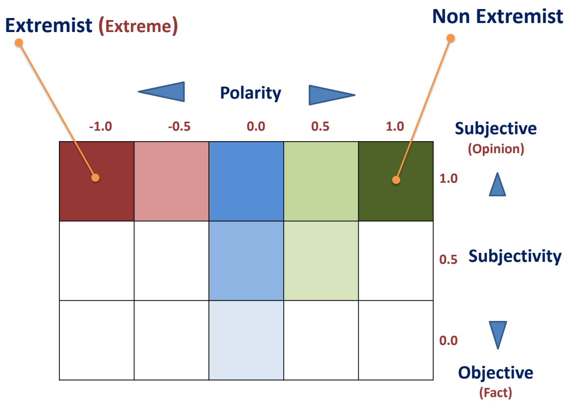

Figure 5.5: Subjectivity and Polarity axes [1].

taking into account word order.

A survey (2015) showed that (83%) term weighting and text-based suggestion systems in digital libraries use TF-IDF [117; 118]. This technology converts text into numbers with mathematical calculations between words and documents. The way TF-IDF works descended from the bag of words model is a numerical statistic intended to reflect how important a word is to a text in a group of texts. The value of TF-IDF increases proportionally to the number of times the word appears in the document and is offset by the number of documents in the group containing the word; Its mathematical equations are given as follows:

$$tf(t,d) = \frac{f_d(t)}{max_{w \in d} f_d(w)} \tag{5.1}$$

$$idf(t,D) = ln(\frac{|D|}{|\{d \in D : t \in d\}|} \tag{5.2}$$

$$tf - idf(t,d,D) = tf(t,d) * idf(t,D) \tag{5.3}$$

Where:

- fd(t): Frequency of term t on document d.

- D : Corpus of documents.

Example Using TF & TF-IDF: (Doc 1 = "The car is driven on the road"; Doc 2 = "The truck is driven on the highway")

Table 5.2: Calculating TF and TF-IDF for corpus doc.

Word	TF	TF	IDF	TF*IDF	
	Doc 1	Doc 2		Doc 1	Doc 2
The	1/7	1/7	log(2/2) = 0	0	0
car	1/7	0	log(2/1) = 0.3	0.043	0
truck	0	1/7	log(2/1) = 0.3	0	0.043
is	1/7	1/7	log(2/2) = 0	0	0
driven	1/7	1/7	log(2/2) = 0	0	0
on	1/7	1/7	log(2/2) = 0	0	0
the	1/7	1/7	log(2/2) = 0	0	0
road	1/7	0	log(2/1) = 0.3	0.043	0
highway	0	1/7	log(2/1) = 0.3	0	0.043

We will synthesize the sentiment analysis approach with the TF, and TF-IDF approaches to converting a dataset textual into a vector of numeric data (Vectorization) [46; 119]. To be ready to learn and test machine learning classifiers.

5.2.2.3 Create dictionaries of features

A dictionary of features is a set of words and sentences that a workbook can use to evaluate the classification of the text to obtain a more accurate classification and reduce the margin of error [120; 121]. The use of a dictionary of features and the presence of a dataset with multiple classifications (more than three) provides several tests on the dataset focusing on a particular species in it or several varieties.

For example: A dataset for five types of birds in which five categories can be created other data sets, including - Birds of prey and non-rapid birds or birds that fly and birds that do not fly), and the results of the measurements are different. The user chooses the model that provides the best results with the classification he wants, and in all cases, the dictionaries of features remain constant.

5.2.3 Used classifiers

Adopting a group of sixteen machine learning classifiers is the most common classifier in much of the literature to perform the classification process and control the characteristics of each algorithm to reach the best results. The following four metrics will be the decisive factor in making the right decision in the results of the machine learning group.

Table 5.3: Machine Learning Classifier

Nu	Abbr	Classifier	Nu	Abbr	Classifier
01	DTC	Decision Tree Classifier	09	BagC	Bagging Classifier
02	RFC	Random Forest Classifier	10	LSVC	Linear SVC
03	ETC	Extra Trees Classifier	11	GNB	Gaussian Naive Bayes
04	MLPC	MLP Classifier	12	BNB	Bernoulli Naive Bayes
05	DC	Dummy Classifier	13	MulNB	Multinomial Naive Bayes
06	ABC	AdaBoost Classifier	14	RC	Ridge Classifier
07	SGDC	SGD Classifier	15	P-AggC	P-Aggressive Classifier
08	G-BC	G-Boosting Classifier	16	K-NC	K-Neighbors Classifier

5.2.4 Evaluations criteria

The dataset in its numerical form in the vectors generated by the proposed NLP technique is passed to the set of linear intelligent classifiers. To get a good rating for a machine learning classifier and to give us good rating accuracy when evaluating, we use the following four criteria:

5.2.4.1 Accuracy:

A metric that gives the classifier the accuracy of the training or testing on the dataset. ISO 5725-1 [122] accuracy is a general term to describe how close a measurement is to the actual value (true value). Applying the term to a dataset of a similar nature includes random and systematic error components, which are calculated as follows :

$$Accuracy = \frac{truepositives + truenegatives}{totalexample} \tag{5.4}$$

5.2.4.2 F1_score:

They evaluate information and many machine learning models, particularly in natural language processing. The F1-score means that the model is perfect [123] and calculated as follows: It is the harmonic mean of the precision and recall.

$$F1 - score = \frac{2}{\frac{1}{recall} \times \frac{1}{precision}} = 2 \times \frac{precision \times recall}{precision + recall} \tag{5.5}$$

Where

$$Precision = \frac{truepositives}{truepositives + falsepositives} \tag{5.6}$$

And

$$Recall = \frac{truepositives}{truepositives + falsenegatives} \tag{5.7}$$

5.2.4.3 Mean_Squared_Error (MSE)

A function calculates the squared error and is the measure of risk corresponding to the expected value of the error or squared loss [124]. It is also a measure of the quality of the estimator and is always positive; values closer to zero are the best [125], and the mean squared error (MSE) estimated over is defined as:

$$\textcolor{red}{MSE} = \frac{1}{q}\sum_{i=n+1}^{n+q}(Y_i - \widehat{Y}_i)^2 \tag{5.8}$$

5.2.4.4 Confusion matrix

It is a matrix(N x N) of the number of categories of the dataset in which the classification model has been tested. It is used to evaluate the performance of this classification model and represents the number of target groups. The actual values of (N) are compared with those predicted by the machine learning classifier model. [126; 127].

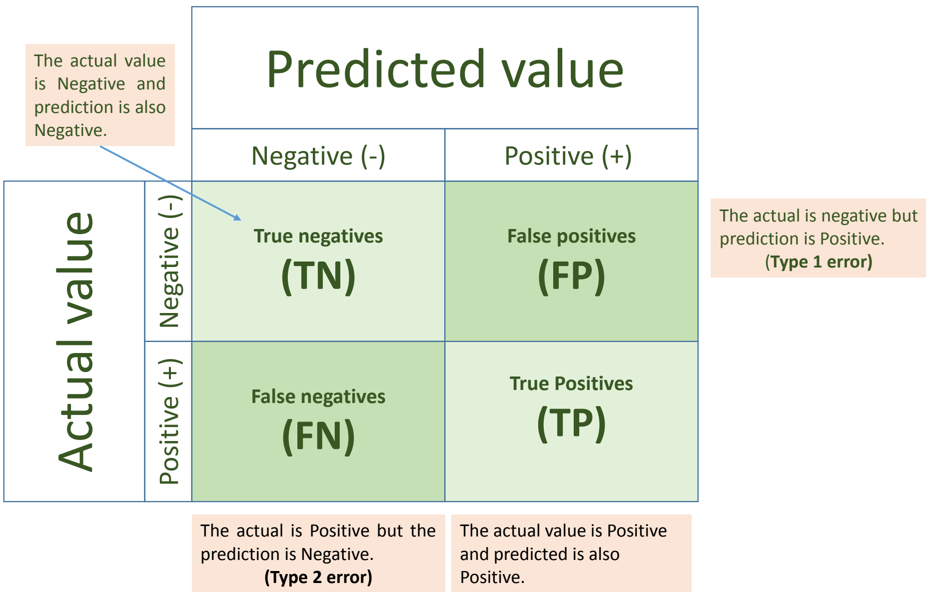

Figure 5.6: Confusion Matrix for Machine Learning
Binary classification problem (N=2)

It is the number of correct and incorrect predictions to evaluate the classifier according to each category in matrix form. When making predictions, it compares the actual target values of the prediction of the machine learning classifier. It gives insight into the problems and errors made by the classifier and identifies the types of overlapping errors [128].

5.3 Conclusion

In this chapter, the structure of the general architecture of the thesis is detailed. The details focused on the most basic three levels in it :

1. The level of pre-processing of the first dataset received from social networks.
2. The second level is retraining the data set on natural language processing techniques to extract its features through dictionaries or numerical data.
3. The level The third is training and testing a set of machine learning classifiers.

Finally, collecting the results of the proposed measures as the study showed the possibility of improving the results of the metrics during the training and testing process.

The following is a presentation of this approach at three levels for a proposed dataset, through which we try to:

1. Detect the content of the extremist text.
2. Classify the types of extremist texts.
3. Then, classify the extremist religious text.

Finally, indicate the percentage of improvement in the metrics results, and discuss and detail the applied approach.

Chapter 6

Experimental Results and discussions

6.1 Introduction

During the previous chapter, we presented the relationships between the basic levels of the study in the form of purely theoretical approaches, from the beginning of collecting the raw data set from social networks, passing through three levels (the proposed data set, the proposed natural language processing techniques and the proposed machine learning classifiers) until obtaining the trainer model. It can detect and classify extremism and its types through new data from social networks.

In this chapter, and after relying on two composite sets of raw text data from the original social networks (Twitter, Facebook), each representing the opinion of its editor, we apply the general architecture of the proposed approach to them.

It exploits data graphs, tables, and results of the four measures across the three stages of training and testing of the proposed machine learning classifiers to visualize results and their improvements. During the attic of the application, all the results of the study obtained are analyzed and discussed in detail, taking into account all the details of the stages of the three levels and all possible cases of the results. A complete discussion of the results and the different approaches and methods led to them, with interpretation and evaluation of the results and comparison with those in various other works of literature to give them credibility.

6.2 Apply the proposed approaches

Extremism is not necessarily caused by religious texts, political texts, intellectual texts, or other social texts that can be circulated on social networks. Therefore, the concepts of social relations (religion, politics, thought, and transactions of all kinds) must be separated from the concept of extremism. Since the relationship between social texts in all their forms and extremist texts is not deterministic, our approach will depend on separating them in terms of methodology, results, and discussion. In the general curriculum of the study, we propose three independent approaches in terms of their establishment. Each approach is established on its own and sequentially in terms of using the resulting compilations to detect extremism, classify its type, and then classify the nature of religion for the proposed text, as shown in the following paragraphs:

- Detect extreme text (extremist text, non-extremist text).
- Extreme text classification (religious extremism, political extremism and intellectual extremism).
- Classification of extremist religious text (Islamic religious extremism, Christian religious extremism, Jewish religious extremism, atheistic extremism).

From the above, the result is three models. The first model works to detect extremism in the text (extremist text, non-extremist text). The second model works to classify the extremist text according to the proposed types (religious extremism, political extremism, and intellectual extremism). The third model classifies the type of extremist religious text that occurs (Islamic extremist text, Christian extremist text, Jewish extremist text, atheistic extremist text, normal text).

The following is a detailed explanation of the results of the stages of our approach in Three steps: collect and organize the data set, extract the features and features of the data set using the proposed natural language processing technology, and train and test the proposed set of machine learning classifiers.

The following is a detailed explanation of the results of applying the proposed approach at three levels:

- The first level is an intelligent model that detects extremist and non-extremist text.
- The second level is an intelligent model that classifies the extremist text resulting from the first model to which type of extremism it belongs (religious extremist text, extremist political text, ideological extremist text).
- The third level is an intelligent model that classifies the extremist religious text resulting from the second model for any religion (Islamic extremist text, Christian extremist text, Jewish extremist text, atheistic extremist text, normal text).

During our study, we use supervised Machine Learning to beat the databases and to get the best results; we divide the data set into two parts, (80%) for training and (20%) for testing [129].

6.2.1 Detecting and classifying extremism

Our work focuses on detecting and classifying extremism that negatively impacts community life. We found that the three types of extremism (Intellectual, political, and religious) can ultimately lead to violent extremism. This extremism influences individuals and groups to carry out heinous acts.

In the next section, the dataset's extreme detection and classification approaches, natural language processing technology, and machine learning classifiers differ in classifying extremist texts. According to sentiment analysis techniques, the extremist detection approach classifies texts into two categories (extreme text and non-extremist text). The extremist text typology approach classifies text into three categories (religious, political, and intellectual) according to the source of the data set.

6.2.1.1 Data collection

The dataset is divided in terms of its source into two parts; the first source is the website kaggle.com, where we found the most significant part of the dataset on religious extremism (Pro-ISI). This data set consists of 17,000 tweets from more than 100 Pro-ISIS fans. It aims to be a counterweight to how Isis uses the Twitter dataset. Including most of the terms considered essential features of the study topic. Without further editing or specification (isis, ISIL, Daesh, Islamic State, Raqqa, Mosul, Islamic state... etc.).

The second source for the rest of the data group is the Twitter developer; it is the remainder of the dataset and it is for the organizations, individuals, political extremists, and ideologues that have the most influence on individuals and peoples, e.g.(@ProudBoysUS Proud Boys (lit. "The proud boys") is an American neo-fascist organization. In the USA, accepting only men among its members promotes and is involved in acts of political violence in the United States).

Our work is to train and test the set of machine learning on our dataset in two stages. The first stage: is a classifier that categorizes tweets according to sentiment (extremist tweet, non-extremist tweet). Second stage: a classifier that categorizes tweets based on the nature of the tweet (intellectual tweet, political tweet, religious tweet):

a) - We distinguish between extremist and not extremist tweets by using a sentiment analysis approach for all tweets as shown in Table 6.1.

Table 6.1: Number of tweets by class.

Class of extremism	Tweets extreme	Tweets Not extreme	Total
Number of tweets	5612	23811	29423

b) - Table 6.2 determines the type of extremism each tweet belongs to, we divided the dataset into three classes (Religious extremism, Political extremism, and Intellectual extremism).

Table 6.2: Number of Tweets by nature of extremism.

Class of extremism	Religious	Political	Intellectual	Total
Number of tweets	20627	5348	3848	29423

c) - In both cases, the tweet features were extracted from the dataset using one of the NLP techniques Term Frequency, Inverse of Document Frequency (TF-IDF).

6.2.1.2 Data cleaning

The first step in analyzing a dataset is cleaning it to ensure it is free of inaccurate, corrupt, and redundant information so that it is ready to extract its features once the cleaning is complete. The main advantages of data cleaning are improved decision-making and cost-effectiveness [130]. This work will explain how to clean up the dataset (tweets) to prepare it.

To highlight and show the features of extremism in the dataset, we purify it in two stages: after collecting tweets in a (CSV) file and depending on the source of the tweets, we add one column according to the type of extremism (religious, political, and intellectual). In this step, as presented in Figure 5.5, we analyze the word cloud of the raw dataset in terms of the high prominence of the most important words, showing.

- In the first step, passing the raw dataset Figure (6.1) over the dataset's usual cleanup function, see Figure 6.2), which focuses on removing stop words, punctuation, hyperlinks, and numbers to get a partially clean dataset.

- In the second stage, delete the padding words (Function 1) that are not related to the subject of the study (incorrect, irrelevant, or incomplete words) and have a high frequency. Table (6.3) presents how to remove these words Figure (6.3).

Algorithm 1 Delete Padding Words

```
Data: Dataset [ ]
Data: Dictionary-Padding-Words [ ]
Result: New-Dataset # Without Padding Words
Initialization
 while Not at end of Dictionary-Padding-Words do
   Read word-PW of Dictionary-Padding-Words
    while not at end of Dataset do
      Read word-D of Dataset
       if word-PW.Dictionary-Padding-Words == word-D.Dataset then
         DELETE word-D FROM Dataset
      else
         GoTo NEXT
      end
   end
   New-Dataset = Dataset;
    RETURN New-Dataset # Without Padding Words
end
```

To this end, we created a dictionary that we named (Padding Words Dictionary) in Table (6.3). We created a function (Function: Delete Padding-Words (1)) that scans the dataset with the padding words dictionary and removes it so that the result appears in the word cloud as shown in Figure (6.3).

After implementing the function of removing the padding words (Function (1)), an apparent result appears in the word cloud, as presented in Figure (6.3), where words related to the topic are visible, indicating that they are in the foreground in terms of their frequency value most common in the dataset. Hence, the dataset is purer and ready for the following stages.

Table 6.3: Dictionary of padding words with the high frequency of each word in the dataset.

Padding words	['amp' : 889, 'today' : 826, 'people' : 767, 'new' : 609, 'near' 576, 'https...' : 597 , 'un' : 524, 'time' : 480, 'news' : 336, 'years' : 306 , 'day' : 301, 'report' : 294, 'want' : 287, 'twitter' : 243, , 'days' : 240, 'dr' : 236, 'no' : 236, 'fsa' : 232 , 'call' : 226, 'think' : 196, 'http...' : 153 , 'year' : 152, 'https' : 145, 'look' : 138, 'account' : 125, 'wj' : 110, 'thank' : 106, 'h...' : 100]

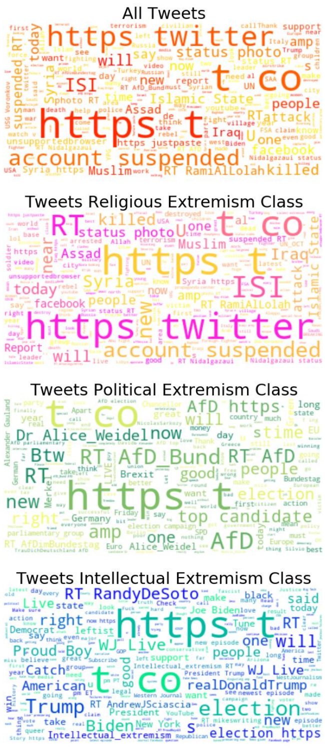

Figure 6.1: The raw dataset.

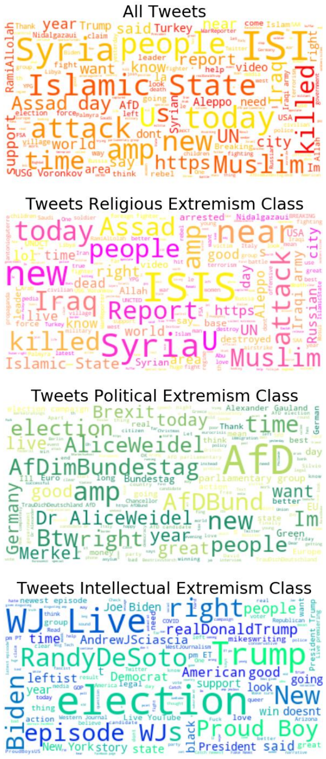

Figure 6.2: Dataset cleaning.

Figure 6.3: Dataset cleaning without padding words.

6.2.1.3 Feature extraction

Words and expressions that include actions, ideas, characters' names, references, books, and articles in the extremism dataset are among the essential textual features used to discover and classify extremism [131]. Through this, it is possible to discover the existence of extremism or its absence and then classify it (Religious, Political, Intellectual) 6.2. In order to extract these features accurately, we use within our approach two techniques used together to achieve a better result:

6.2.1.4 Sentiment analysis

Sentiment analysis, also called opinion mining, is an application for natural language processing and computational linguistics of text analysis to discover and categorize personal opinions (for example, Facebook comments or Twitter Tweets) [132]. Sentiment analysis aims to identify the position or emotional state of the author when writing or to convey the emotional impact the author wishes to have on the reader [115]. Usually, objective tweets are without feelings, and subjective tweets tend to have positive or negative feelings. The situation or emotional state is determined by classifying the subjectivity of the tweets and then classifying them into positive, natural, and negative tweets. Figure 5.5 principle is used for revealing the subjectivity and polarity of the tweets.

Subjectivity analysis To deduce the features of the most useful Tweets and the most subjective texts that express their editor's opinion. The subjectivity function is applied to the Twitter [133] dataset (6.1) using the extreme text attribute. The subjectivity score of all tweets is calculated to find the subjectivity of all tweets.

Polarity analysis The polarity analysis function is applied to the dataset (6.1) to find out how negative and positive the tweet is [134]. Sentiments are averaged across all tweets to find the accuracy of sentiment analysis and tested based on the class of extremes trait.

We can calculate the negative or positive tweets through these two sentiment analysis features. We rely on the TextBlob package in Python as a convenient way to do many NLP tasks, mainly calculating polarity and subjectivity from which to infer tweet sentiment.

For example, The results obtained by applying TextBlob on the tweet are

Table 6.4: Example of sentiment analysis of a tweet after cleaning it.

Tweet	Tweet clean	polarity	subject	Sent
RT @*RamiAlLolah*: #Syria—n revolution actishits are unhappy about #ISIS gains over #Folan thugs in #DeirEzzor. Donkeys will never *learn*..	ramiallolah syrian revolution actishits unhappy is gain Folan thug director donkey learn	-0.6	0.9	-1

shown in Table (6.4). These results indicate that it has a polarity of about (-0.6), which means that it is somewhat negative, and approximately (0.9), which is largely subjective. The tweet sentiment becomes negative by applying the get-Analysis function to the resulting value[135].

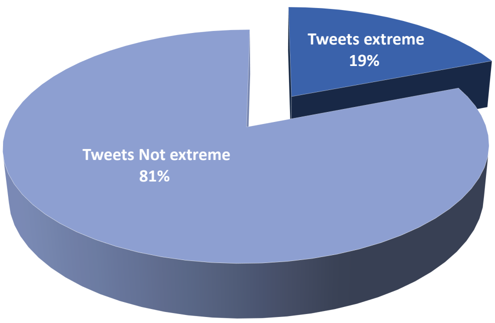

Figure 6.4: Distribution of extreme texts classes.

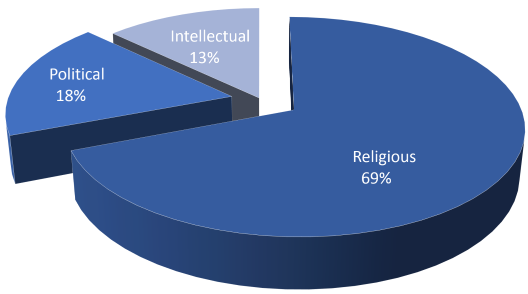

Figure 6.5: Distribution of extreme text types.

The product of the relative circle numbers Figure(6.4) shows that the Distribution of feelings after calculating the subjectivity and polarity of all tweets in the data set shows that the proportion of extreme tweets is much lower than that of non-extremists, which is inherently unbalanced,

The results of the relative circle numbers Figure(6.5) show that the distribution of the categories of the types of extremism in the data set is unbalanced, the religiously extremist tweets are a large number due to the other two categories, and this is due to their source. The Religious Extremist Tweet Dataset is a resource ( ) that provides ready-made data sets with a vast pilgrimage. The data set of politically and ideologically extremist tweets was collected from the Twitter application, it was manually organized to be ready to work on, and this takes significant time and effort.

6.2.1.5 TF-IDF NLP technique (Term Frequency, Inverse of Document Frequency)

Statistical supervised linear classifiers use mathematics to train machine learning classifiers. We must convert text to numbers to make linear classifiers work with text. This section will discuss the TF-IDF approach that descends from the Bag-of-Words Model.

The TF-IDF technique is the text back-frequency of the short-term frequency, a numerical statistic that aims to reflect how important a word is to a text in a set of texts. A survey (2015) showed that (83%) of term weighting systems and text-based suggestion systems in digital libraries use TF-IDF [136].

The weighting factor is often used in information retrieval searches, text mining, and user modeling. The TF-IDF value increases proportionally to the number of times a word appears in the document. It is offset by the number of documents in the corpus that contain the word, which helps to adjust that some words appear more frequently in general. To test our sentiment analysis model, we will synthesize the sentiment analysis approach with the TF-IDF approach to convert textual data into digital data that a Machine Learning classifier can use.

We relied on the Python Scikit-Learn library [137] with a TF-IDF-vectorizer class that can convert text features into TF-IDF feature vectors. The idea behind using the TF-IDF approach [138] is that words that appear less in all tweets and more in one tweet contribute more to ranking, with each word having the same weight in all tweets. We convert raw data into a sparse n-gram feature matrix to reach this end.

The results of the four metrics (5.4, 5.5, 5.8 and 5.6) on training and testing of the proposed set of (16) classifiers (5.3) for the three levels (6.5)5of the dataset for extremism detection and classification are discussed. The last metric: the confusion matrix (5.6), is used to determine the accuracy of the uncertainty in the first and second models.

Table 6.5: Levels of training and testing.

Level	Operation
L1	Machine learning on the raw dataset
L2	Machine learning on the clean dataset
L3	Machine learning on the clean dataset without padding words

In the following, we will present the results and discussions of the application of the three levels of the approach within the following two classifications.

NB : The results values will be three numbers before the break to facilitate comparison and analysis. This is due to the convergence of the results.

6.2.2 Extremism detection Results and discussions

This work is a comparative study of the results of applying sixteen (16) classifiers of Scikit-Learn Machine Learning (5.3) to create an extremist text detection model that can perform the same process differently [137].

6.2.2.1 Extremism detection: Results

Table (6.6) shows the training and testing results of a group of sixteen (16) supervised classifiers (5.3) on the extreme text detection data set (6.9)).

Table 6.6: The results of the first model (Extremism detection).

Metrics	Training Accuracy			Testing Accuracy			F1-score			MSE		
Classifiers	L1	L2	L3	L1	L2	L3	L1	L2	L3	L1	L2	L3
1-DTC	0,996	0,995	0,999	0,812	0,885	0,961	0,787	0,857	0,935	0,337	0,197	0,069
2-RFC	0,996	0,995	0,999	0,842	0,912	0,967	0,817	0,888	0,944	0,275	0,148	0,059
3-ETC	0,996	0,995	0,999	0,859	0,914	0,969	0,836	0,891	0,949	0,236	0,144	0,053
4-MLPC	0,905	0,916	0,959	0,838	0,869	0,921	0,816	0,828	0,849	0,280	0,235	0,171
5-DC	0,523	0,597	0,721	0,507	0,584	0,707	0,224	0,246	0,276	0,493	0,416	0,293
6-ABC	0,630	0,707	0,800	0,610	0,693	0,791	0,475	0,542	0,584	0,436	0,339	0,217
7-SGDC	0,847	0,895	0,965	0,822	0,878	0,952	0,790	0,842	0,922	0,259	0,172	0,067
8-G-BC	0,607	0,675	0,767	0,590	0,662	0,755	0,435	0,471	0,459	0,442	0,360	0,250
9-BagC	0,979	0,982	0,995	0,851	0,903	0,964	0,827	0,877	0,940	0,267	0,165	0,061
10-LSVC	0,890	0,925	0,978	0,861	0,914	0,970	0,840	0,891	0,952	0,230	0,132	0,044
11-GNB	0,590	0,492	0,338	0,524	0,442	0,305	0,523	0,453	0,325	0,972	0,977	0,959
12-BNB	0,771	0,852	0,937	0,723	0,824	0,904	0,692	0,780	0,851	0,534	0,316	0,167
13-MulNB	0,806	0,849	0,879	0,761	0,815	0,842	0,709	0,752	0,697	0,411	0,261	0,183
14-RC	0,878	0,917	0,972	0,851	0,899	0,958	0,827	0,872	0,932	0,240	0,150	0,062
15-P-AggC	0,895	0,917	0,973	0,836	0,877	0,947	0,818	0,853	0,916	0,272	0,195	0,066
16-K-NC	0,688	0,738	0,806	0,551	0,629	0,744	0,389	0,427	0,450	0,504	0,403	0,262

6.2.2.2 Extremism detection: Discussions

Table 6.6 shows the results of the three metrics defined in (5.4, 5.5 and 5.8) during the three training and testing levels illustrated in Table (6.5) for the group of (16) machine learning classifiers (5.3). To analyze and discuss the results of the scales for the various levels, we rely on cylindrical bar graphs to give a figure showing the percentages of the measures. We use the confusion matrix to discuss the optimization of the accuracy result of the resulting classifier with the dataset.

The cylindrical bar graph of Figure (6.6) and Table (6.6) shows the training and testing of the set of machine learning classifiers on the raw dataset. The training accuracy results for the classifiers (1, 2, 3, 4, 9, and 10) exceeded 90%, which are acceptable results. It is also shown that the test accuracy results for classifiers (1, 2, 3, 4, 7, 9, 10, 14, and 15) did not exceed 86.1% which are close to acceptable results. The MSE results for all classifiers were higher than 20%,

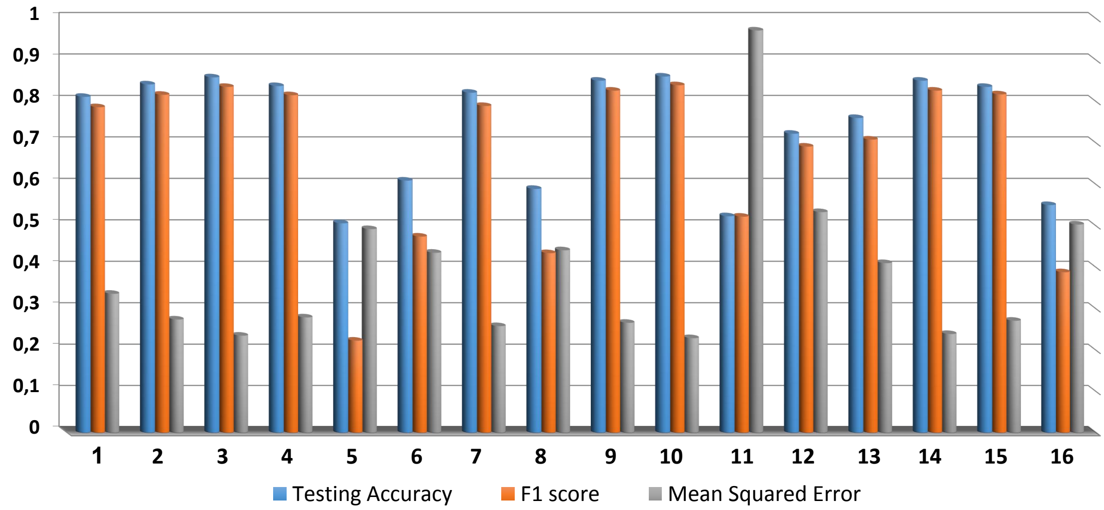

Figure 6.6: ML output using the raw dataset (level L1 : for extremism detection).

which disturbs the accuracy values of the classifiers. At this level L1, the linear SVC classifier scored the best test accuracy on the raw dataset 86.1% and the best results for the other metrics as shown in Table (6.7), but the results are still subject to improvement.

Table 6.7: The best results of the model in level L1 (ML on the raw dataset- for extremism detection).

Accuracy metric	Best classifier	Accuracy
Training Accuracy	Decision Tree Classifier	99,6%
Testing Accuracy	Linear SVC	86,1%
F1 score	Linear SVC	84,0%
Mean squared error	Linear SVC	23,0%

The cylindrical bar graph of Figure 6.7 shows the training and testing of the set of machine learning classifiers on the clean dataset (Level 2, see Table 6.8), and it is clear that the training accuracy results of the classifiers have improved in most classifiers and are acceptable for training. It is also shown that the test accuracy results for classifiers may exceed 90%, which is satisfactory and acceptable. The results in Table 6.6, Level 2 show that the MSE values of most

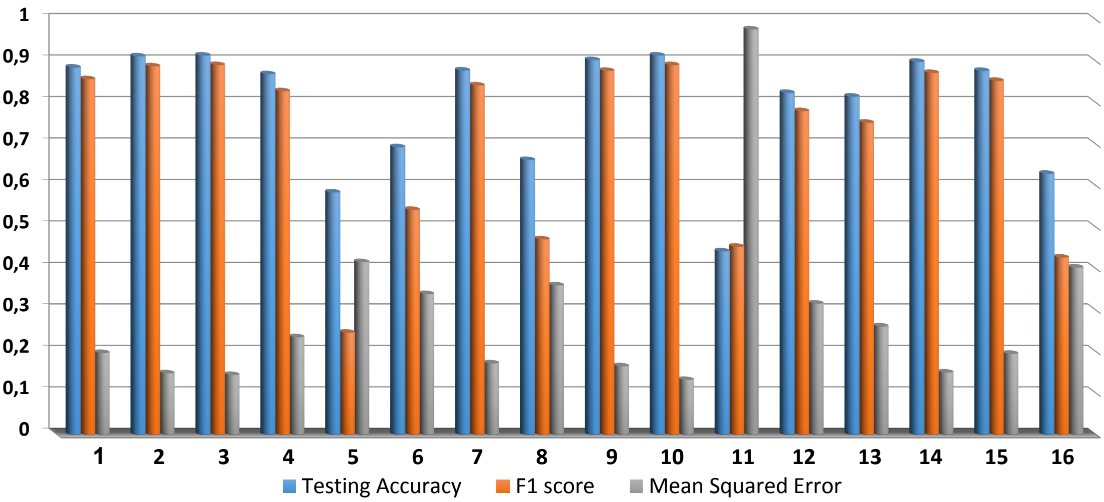

Figure 6.7: ML output on the clean dataset (level L2 :for extremism detection).

classifiers became less than 20%, which gives credibility to the accuracy values of the test classifiers.

Table 6.8: The best results of the model in level L2 (ML on the clean dataset-for extremism detection).

Accuracy metric	Best classifier	Accuracy
Training Accuracy	Decision Tree Classifier	99,5%
Testing Accuracy	Random Forest and Linear SVC	91,4%
F1 score	Linear SVC	89,1 %
Mean squared error	Linear SVC	13,2 %

Classification N: 11 with Gaussian Naive Bayes did not improve but worsened. At this level L2, the Decision Tree Classifier scored the best training accuracy, Random Forest Classifier and Linear SVC classifiers scored the best test accuracy, and they scored the best F1_score and MSE values shown in Table 6.8. The SVC linear classifier can be adopted as the best classifier for this level with an improvement rate in the measurement results with a testing accuracy of $5,3\%$, F1_score of $5,1\%$ and MSE of $-9,8\%$.

From Figure 6.8, it can be seen that most of the classifiers have exceeded the

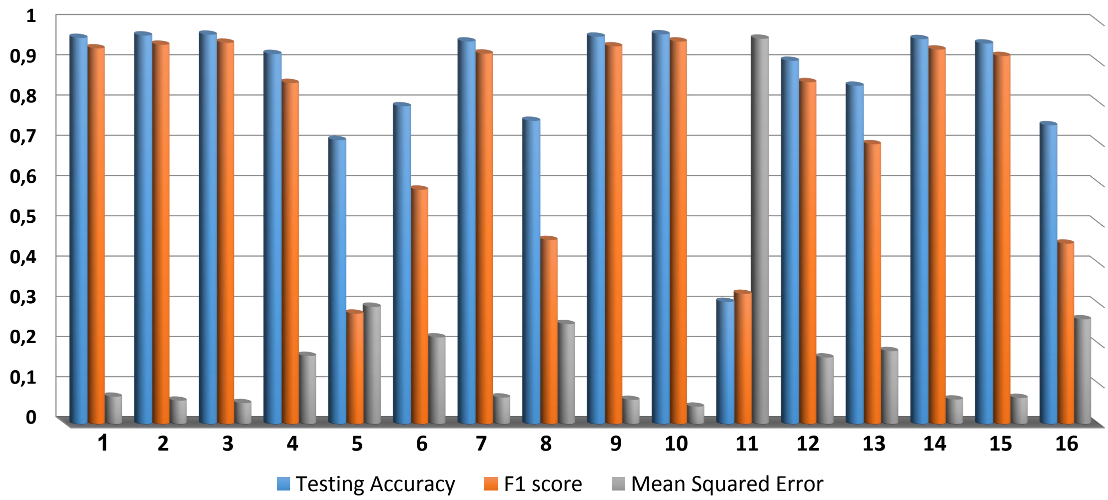

Figure 6.8: ML outputs on the clean dataset without padding words (level L3: for extremism detection).

threshold of 90% in the test accuracy results, of which six have exceeded 95%. The result in Table 6.9 shows that the linear support vector classifier gave the best result in the test accuracy of 97%. The results presented in Table 6.6 show that the average MSE values of the classifiers(1, 2, 3, 7, 9, 10, 14 and 15) to less than 20%, which are satisfactory values and give strong credibility to the accuracy of the test.

Table 6.9: The best results of the model in level L3 (ML on the clean dataset without padding words - for extremism detection).

Accuracy metric	Best classifier	Accuracy
Training Accuracy	Decision Tree Classifier	99,9%
Testing Accuracy	Linear SVC	97,0%
F1 score	Linear SVC	95,3%
Mean squared error	Linear SVC	4,40%

In general, the results of Level 3, as presented in Table (6.9) indicate that the function for cleaning tweets from high-frequency padding words works efficiently with a majority of classifiers, as shown in Figure (6.8) except for the Gaussian

Naive Bayes, which confirms the necessity of this function in improving the metrics results. This offers an effective model for classifying extremist tweets. The improvement from Level 1 to Level 3 was estimated as a percentage as follows: training accuracy of 0.3%, testing accuracy of (10.9%), F1_score of (11.3%) and MSE of (-18.6%).

Table 6.10: Percentage of improvement for the second level (L2).

Metrics	**Percentage**
Training Accuracy	0.3%
Testing Accuracy	$10,9\%$
F1 score	$11,3\%$
Mean squared error	$-18,6\%$

6.2.3 Extremism classification Results and discussions

This work is a comparative study of the results of applying sixteen (16) Scikit-Learn Machine Learning classifiers (5.3), to create an extreme text type classification model that can perform the same process differently [137].

6.2.3.1 Extremism classification: Results

Table (6.11) shows the training and testing results of a set of sixteen (16) moderated classifiers (5.3) on the Extreme Text Types Classification dataset (6.11)).

Table 6.11: The results of the second model (for extremism classification)..

Metrics	Train Accuracy			Test Accuracy			F1-score			MSE		
Classifiers	L1	L2	L3	L1	L2	L3	L1	L2	L3	L1	L2	L3
1-DTC	0,988	0,989	0,992	0,852	0,853	0,890	0,771	0,777	0,826	0,303	0,297	0,195
2-RFC	0,988	0,989	0,992	0,885	0,887	0,927	0,813	0,819	0,880	0,236	0,230	0,128
3-ETC	0,988	0,989	0,992	0,889	0,890	0,921	0,821	0,822	0,875	0,227	0,221	0,146
4-MLPC	0,919	0,918	0,971	0,872	0,874	0,929	0,764	0,761	0,878	0,288	0,290	0,122
5-DC	0,704	0,704	0,704	0,696	0,696	0,696	0,274	0,274	0,274	0,643	0,643	0,643
6-ABC	0,766	0,769	0,780	0,753	0,758	0,770	0,506	0,523	0,603	0,505	0,486	0,459
7-SGDC	0,925	0,919	0,955	0,906	0,901	0,936	0,842	0,835	0,894	0,204	0,220	0,111
8-G-BC	0,738	0,738	0,748	0,728	0,728	0,738	0,378	0,377	0,431	0,607	0,608	0,571
9-BagC	0,981	0,982	0,986	0,870	0,872	0,907	0,792	0,797	0,852	0,268	0,253	0,170
10-LSVC	0,955	0,956	0,971	0,906	0,907	0,927	0,843	0,847	0,877	0,199	0,194	0,136
11-GNB	0,802	0,802	0,855	0,781	0,782	0,823	0,712	0,712	0,749	0,502	0,502	0,395
12-BNB	0,916	0,917	0,909	0,903	0,903	0,898	0,846	0,845	0,855	0,202	0,202	0,166
13-MulNB	0,917	0,917	0,929	0,902	0,903	0,916	0,836	0,836	0,857	0,218	0,214	0,171
14-RC	0,929	0,929	0,949	0,902	0,902	0,922	0,837	0,838	0,870	0,208	0,208	0,152
15-P-AggC	0,968	0,970	0,980	0,897	0,895	0,920	0,830	0,833	0,868	0,212	0,213	0,153
16-K-NC	0,817	0,815	0,825	0,758	0,756	0,769	0,534	0,528	0,566	0,497	0,500	0,468

6.2.3.2 Extremism classification: Discussions

Table 6.11 shows the results of the three measures during the three training and testing levels (L1, L2, and L3(6.6)) of the machine learning set. To analyze the results of Table 6.12, we rely on the following bar graphs:

From the cylindrical bar graph in Figure 6.9, which represents the results of the metrics using the machine learning set in the raw dataset, it is clear that the training accuracy for the (11) classifiers exceeded 90%, which is acceptable results. Only five classifiers exceeded the generally accepted test accuracy of 90%. Tables (6.11 and 6.20) record Linear SVM has the best MSE value of 19.9%. The Decision Tree Classifier gave the best training accuracy on the dataset (98.8%), the linear support vector classification classifier gave the best test accuracy on the dataset (90.6%), and the Bernoulli Naive Bayes classifier gave the best F1_score value (84.6%).

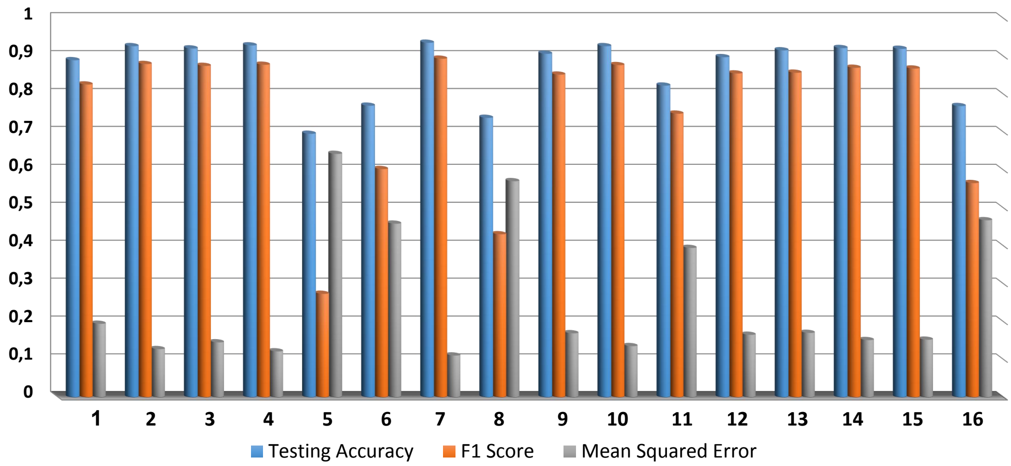

Figure 6.9: The best results of the model in level L1 (ML on the raw dataset- for extremism classification).

Table 6.12: The best results of the model in level L2 (ML on clean dataset- for extremism classification).

Accuracy metric	Best classifier	Accuracy
Training Accuracy	Decision Tree Classifier	98,8%
Testing Accuracy	Linear SVC	90,6%
F1 score	Bernoulli Naive Bayes	84,6%
Mean squared error	Linear SVC	19,9%

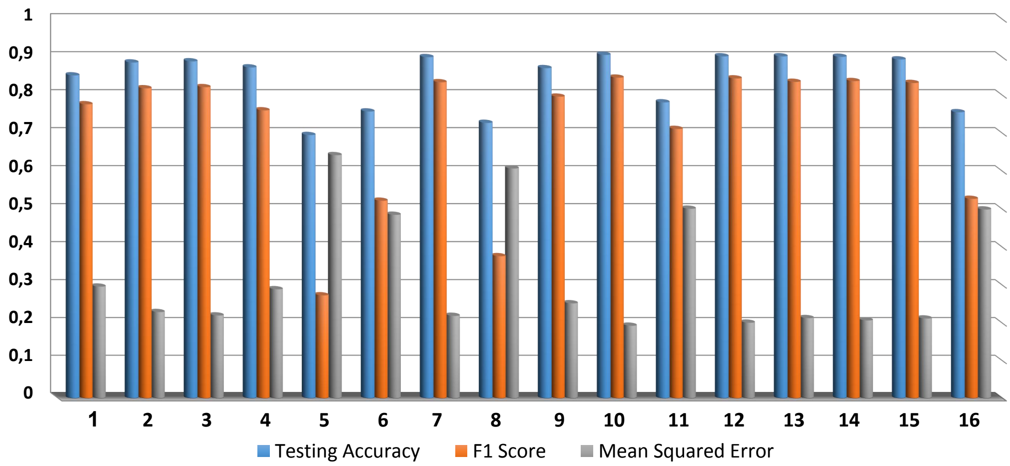

Figure 6.10: ML output using the raw dataset (for extremism classification).

From the cylindrical bar graph of Figure (6.10), which represents the results of the Metrics using machine learning set in the clean dataset Level 2 6.5), it is clear that the training accuracy for the (11) classifiers exceeds 90% with slight improvement during this level, which is acceptable results. Only five classifiers exceeded the generally accepted test accuracy of 90%. From Figure (6.10) and Table (6.13), the Linear SVC has the best MSE value of 19.4%. The Decision Tree classifier gave the best training accuracy of 98.8% on the data set. The linear support vector classification classifier gave the best test accuracy of 90.7% on the dataset and the best F1_score of 84.6%. At this level, the cleaning process with normal cleaning functions resulted in a slight improvement in the metrics in negligible proportions; only the F1_score improved by $0,02\%$.

Table 6.13: The best results of the model in level 2 (ML on the clean dataset- for extremism classification).

Accuracy metric	Best classifier	Accuracy
Training Accuracy	Decision Tree Classifier	98,8 %
Testing Accuracy	Linear SVC	90,7 %
F1 score	Linear SVC	84,4 %
Mean squared error	Linear SVC	19,4 %

From the cylindrical bar graph (6.11), which represents the results of the Metrics using a machine learning set in the clean data set without padding words Level 3 (6.5), it is clear that the training accuracy for the eleven (11) classifiers exceeds 90% with significant improvement during this level, and the results are acceptable. Nine (9) classifiers exceed the test accuracy of 90%; these are very satisfactory results. An MSE was recorded as a minimum value of 11.1%, which gives credibility to the result of the workbook test.

From Figure (6.9), Tables (6.14 and 6.9) the linear SVC gave the best MSE of 19.4%. The Random Forest classifier gave the best training accuracy on the dataset with 98.8%, the linear support vector classifier gave the best test accuracy

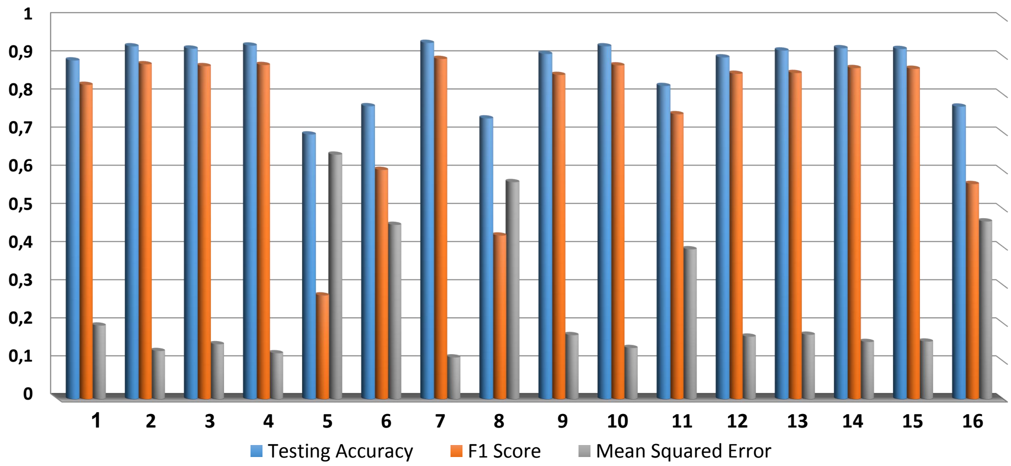

Figure 6.11: ML output on the clean dataset without word padding (for extremism classification).

Table 6.14: The best results of the model in level 3 (ML on the clean dataset without padding word for extremism classification).

Accuracy metric	Best classifier	Accuracy
Training Accuracy	Decision Tree Classifier	99,2 %
Testing Accuracy	SGD Classifier	93,6 %
F1 score	Linear SVC	89,4 %
Mean squared error	Linear SVC	11,1 %

on the dataset with 90.6%, and the best F1_score value of 84.6%. During this level, the cleaning process with the normal cleaning functions with the function to delete the padding words resulted in a significant improvement in the metrics in acceptable percentages as follows: (training accuracy of 0.4%, testing accuracy of 2.6%, F1_score 4.8% and MSE of −8.8%).

Table 6.15: Percentage of improvement for the third level (L3).

Metrics	**Percentage**
Training Accuracy	$0,4\%$
Testing Accuracy	$2,6\%$
F1 score	$4,8\%$
Mean squared error	$-8,8\%$

We note in the Tables (6.6 and 6.11) that the results of the Metrics recorded an improvement from the Level L1 (raw dataset) to Level L3 (6.5). Level L3 cleans the dataset without the padding words with acceptable percentages. Notably, the percentage improvement in detecting extreme text is better than using the same classifiers and the same data set to classify the extreme text. Thus, our approach of combining sentiment analysis and TF-IDF with a padding word cleaning function proved to work best with the extremism detection model. Its results are also very accurate with the model for classifying extremism (religious, political, intellectual), and its performance is acceptable. The high-frequency padding word cleaning function led to high-accuracy results in terms of detection of extremism and classification of extremism, which decision-makers can rely upon by analyzing social texts.

6.2.4 Classification of extremist religious text

This approach aims to classify extremist religious texts into five classifications of religious texts (Christian, Jewish, Muslim, atheist, or normal text). This classification aims to classify extremist religious texts through the texts of users of social networks.

Before using classification models, data must be gathered and pre-processed. To improve the classification accuracy, incomplete, inconsistent, and noisy real-world data needs to be pre-processed with some techniques, including missing-data filtering, data integration, data normalization, and feature selection.

6.2.4.1 Data collection

Collecting and classifying data is a rather difficult task because it takes much time, and much knowledge is required to reach a final decision that defines the category of each comment (Christian, Jewish, Atheist, Islamic, or natural).

The latter took us from five to six weeks and two reviewers (the third until the dispute between reviewers is resolved by a majority vote) to collect (3373) comments from different religious groups on Facebook, among them Jewish Spirit, Christian Support Group, Islam Muslim, Global Islamic Orator Group, the Atheists Group, the Atheist Alliance of America ... Etc. It is then classified into five categories: Christianity, Jewish, Atheist, Islamic, and Neutral 6.1.

After the data is collected, it is cleaned to be free from inaccurate, corrupt, and redundant information so that it is ready for analysis once the cleaning is completed to improve classification decisions and cost-effectiveness.

Table 6.16: Number of comments by class.

Classes	Christianity	Judaism	Atheism	Islamic	Neutral	**Total**
Number of comment	643	716	706	701	607	**3373**

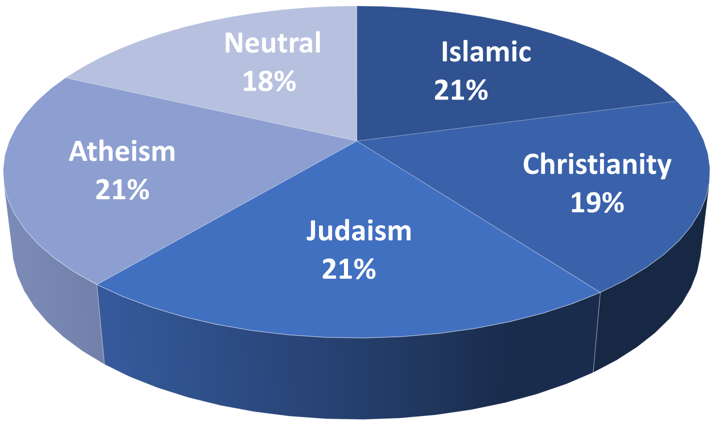

Figure 6.12: Percentages of comments by class

From the table 6.1 and the circle of percentages 6.2 of comments, we notice that the number of comments for the categories is convergent; that is, the data sets are balanced. The following table has an overview of the comments for each category.

We have worked on the data in several different cases in terms of categories

Table 6.17: Examples of classification of some comments.

Classes	Polarity
If God exists, why can't we see Him?	Atheist
If God is fair then why are diseases spread, and why is injustice permitted among people?	Atheist
Subhun Allah Alhamduillah L'ilahaillallahAllahu Akbar.	Muslim
AmeenInshaAllahi will meet my husband in jannahulfirdusAmeen.	Muslim
Amen thank you Jesus my Lord and Savior.	Christian
HALLELUJAH. GLORY TO GOD ALMIGHTY. THANK YOU JESUS CHRIST AMEN.	Christian
just Like The Bible Said Israel Will Turn The Desert Into Paradise Israel Jews,	Jewish
God belongs to the people of the Jews and not to others	Jewish
Please put your menu online!!	Neutral
Wow that really works for me thanks man.	Neutral

and numbers, and in terms of data volume. We will explain what we did on these points:

- Classification into religious and non-religious texts.
- Classification of religious texts.
- Working on balanced and unbalanced categories in terms of the number of data for each category.
- Working on a different number of classifications

6.2.4.2 Feature extraction

The essential characteristic of religious texts is using religious words and phrases and citations from religious books. The traits of each religion have been extracted and compiled into dictionaries.

The primary distinguishing religious texts are using some religious words and phrases and citations from religious books. We have identified the features of each religion that we will explain below:

Features of Christianity text

- Citing from the gospel.
- Citing the names of the Christian God: Holy Spirit, Jesus Christ, Heavenly Father, Lord Jesus, Lord Jesus Christ, the Church, Saviour, Jesus, Jesus, the Scriptures, Christianity, the Messiah, the Gospel, the Spirit... Etc.
- Use of religious terms: Amen, in Jesus Holy Name, Holy Spirit, Thank you, Jesus Christ, Praise you in the highest, Lord Jesus ...etc.
- The use of religious expressions: God the Father, Son Jesus, Holy Spirit, Son of God, the name of Jesus, in Jesus Holy Nam ...etc.

Features of Judaism text

- Citing from the Bible.
- The use of these phrases: glory and eternity for the Lord, Shabat Shalom, God bless Israel, God bless the Jews, I am a Jew, God stands with his children and We are God's chosen people.
- Writing the word of "God" in this way "G-D".
- Citing from rabbis.
- The use of the word "Jehovah", which means the lord.

Features of Atheism text

- Denial of the existence of God.
- Questioning the existence of God.

- Using questioning tools.
- Use of negation tools: I do not believe, do not think, I do not doubt.
- The denial of the divine books and not believing it.
- Atheists use the phrase: I am not religious, I disbelieve in all religions when inventing a god, your alleged god, and morality originates in atheism.
- Citing the most famous name of atheists: Jason Aaron-Clark Adams-larry Adler-Amy Alkon-Robert Altman-Natalie Angier-Aziz Ansari-Isaac Asimov-Scott Atran-Bob Avakian.

Features of Islam text

- Citation of Quranic verses and prophetic sayings.
- Use of religious phrases: Alhamdulilla, Mashaa' Allah, Allahumma Salli WaSallam Wabareek alahabibul LAH sallallahu Alayhy WaSallam, Allah Subhan Wa Taala , La IllahaIlallah... Muhammad Rasulullah... etc.
- Citation of the greatest personalities of Islamic history: Imam Al-Shafi'i , Imam Al-Bukhari .... etc.
- Citation of the names of the Prophets, Messengers and Companions: Muhammad, Adam, Abu Bakr Al-Siddiq, Othman bin Affan ... etc.
- Citing the attributes of Allah (God)
- The use of the Hijri months: Shaaban. Ramadan. Shawwal ... etc

6.2.4.3 Create dictionaries of features

We create our own dictionaries for each religion (Christianity, Judaism, Islam and atheism) based on what distinguishes each religion and on what we found in the comments of the features, which are shown in Table 6.3:

Table 6.18: Number of features by class.

Class	Islamic	Christianity	Judaism	Atheism	**Total**
Number of feature	7389	132	30179	1156	**38856**

Specific features were identified for each religious class and characterized the features of the dataset (3,373 comments), categorized into five classes (Islamic, Christian, Jewish, salt, and neutral). Finally, we create a set of dictionaries for each class.

Table 6.19: Statistics of Christianity dictionary.

Dictionary name	Number of features	Content
D_christ	132	Religious terms from the gospel.
Total	**132**	

Table 6.20: Statistics of Judaism dictionary.

Dictionary name	Number of features	Content
Bible	29158	Religious terms from the Bible
F_Jews	685	The use of the word Chadian, Jehovah, which means Lord
N_famous_jews	194	The famous personalities of Jews history
S_jews	142	Famous sentences of the Jews
Total	**30179**	

6.2.4.4 TF NLP technique (Term Frequency)

Statistical supervised linear classifiers use mathematics to train machine learning classifiers. We must convert text to numbers (0 or 1) to make linear classifiers

Table 6.21: Statistics of Islamic dictionary.

Dictionary name	Number of features	Content
A_Allah	99	The attributes of Allah
D_Muslim	570	Islamic phrases
H_Hadiths	116	Hadiths
H_months	13	Hijri months
N_famous_Islamic	87	The greatest personalities of Islamic history
Py_Muslim	43	The names of the Prophets, Messengers, and Companions
Quran	6347	Quranic verses
S_Quran	114	The names of the surahs of the Qur'an
Total	**7389**	

Table 6.22: Statistics of Atheism dictionary.

Dictionary name	Number of features	Content
A_Atheist	773	Atheists use the phrase: "god of gaps" I am not religious. I disbelieve in all religions; when inventing a god, your alleged god, morality originates in atheism.
N_famous_Atheist	383	The famous personalities of Atheist history
Total	**1156**	

work with text. This section will discuss the TF approaches that descend from the Bag-of-Words Model.

The TF approach is the frequency of a specified term (feature) in the data set after a feature dictionary scan [139]. To determine the frequency, you must:

1. Calculate the times the term appears in the data set.

2. The frequency of the term is adjusted according to the length of the data set (the number of impressions divided by the number of words in the dataset).

3. The logical frequency of the fixer (1 if the term is present, or 0 if the term

does not exist, the data set).

6.2.4.5 Proposed classifications

The diversity of beliefs and religions in the world constitutes a vast field of research and analysis. It classified (232) countries into six main regions: Asia and the Pacific, Europe, Latin America and the Caribbean, North America, Sub-Saharan Africa, and the Middle East and North Africa. This ranking is based on the "Religious Diversity Index" that the researchers created to conduct their studies. According to them, countries are divided into four major groups according to their religious diversity: very high, high, medium, and low [140].

The index is calculated based on the percentage of different groups in a country: believers of major religions (Islam, Christianity, Judaism ...Etc) and non-believers, including atheists [141]. Therefore, we worked on 18 cases; as the following:

Case 1. Two classes: the first class includes Islamic comments only, and the second class includes the rest of the religious and non-religious comments (Christian, Judaism, Atheism, and neutral) and the number of comments in both classes is (C1:700 comments, C2:700 comments).

Case 2. Two religious classes only: the first class includes Christian comments only, and the second class includes the rest of the religious comments (Islamic, Judaism, and Atheism), and the number of comments in both classes is (C1:643 comments, C2:643 comments).

Case 3. Two classes: the first class includes Christian comments only, and the second class includes the rest of the religious and non-religious comments (Islamic, Judaism, Atheism, and neutral) and the number of comments in both classes is (C1:643 comments, C2:643 comments).

Case 4. Two religious classes only: The first class includes only Islamic comments, and the second class includes the rest of the religious comments (Christian, Jewish, and Atheism), and the number of comments in both classes is (C1:700 comments, C2:700 comments).

Case 5. Two religious classes only: the first class includes Atheism comments only, and the second class includes the rest of the religious comments (Islamic, Christian, and Jewish), and the number of comments in both classes is (C1:706 comments, C2:706 comments).

Case 6. Two classes: the first class includes Atheism comments only, and the second class includes the rest of the religious and non-religious comments (Islamic, Christian, Jewish and neutral) and the number of comments in both classes is (C1:706 comments, C2:706 comments).

Case 7. Two classes: the first class includes Jewish comments only, and the second class includes the rest of the religious and non-religious comments (Islamic, Christian, Atheism, and neutral) and the number of comments in both classes is (C1:716comments, C2:716 comments).

Case 8. Two religious classes only: the first class includes Jewish comments only, and the second class includes the rest of the religious comments (Islamic, Christian, and Atheism), and the number of comments in both classes is (C1:716 comments, C2:716 comments).

Case 9. Two classes: the first class includes Islamic comments only, and the second class includes the rest of the religious and non-religious comments (Christian, Judaism, Atheism, and neutral) and the number of comments in each of the first and second classes is(C1:700 comments, C2:2673 comments).

Case 10. Two classes: the first class includes Christian comments only, the second class includes the rest of the religious and non-religious comments (Islamic, Judaism, Atheism, and neutral), and the number of comments in both classes is (C1:643 comments, C2:2730 comments).

Case 11. Two religious classes only: the first class includes Islamic comments only, and the second class includes the rest of the religious comments (Christian, Judaism, and Atheism); the number of comments in both the first and second classes is (C1:700 comments, C2:2065 comments).

Case 12. Two religious classes only: the first class includes Christian comments only, and the second class includes the rest of the religious comments (Islamic, Judaism, and Atheism) and the number of comments in both classes is (C1:643 comments, C2:2122 comments).

Case 13. Two classes: the first class includes Jewish comments only, and the second class includes the rest of the religious and non-religious comments (Islamic, Christian, Atheism, and neutral), and the number of comments in both classes is (C1:716 comments, C2:2657 comments).

Case 14. Two classes: the first class includes Atheism comments only, and the second class includes the rest of the religious and non-religious comments (Islamic, Christian, Jewish and neutral) and the number of comments in both classes is (C1:706 comments, C2:2667 comments).

Case 15. Two religious classes only: the first class includes Jewish comments only, and the second class includes the rest of the religious comments (Islamic, Christian, and Atheism), and the number of comments in both classes is (C1:716 comments, C2:2049 comments).

Case 16. Two religious classes only: the first class includes Atheism comments only, and the second class includes the rest of the religious comments (Islamic, Christian, and Jewish), and the number of comments in both classes is (C1:706 comments, C2:2059 comments).

Case 17. Five classes (Islamic, Christian, Jewish, Atheism, neutral), and the number of comments in each class is (C1:706 comments, C2:700 comments, C3:643 comments, C4:716 comments, C5:608 comments).

Case 18. Four classes: Islamic, Christian, Jewish, and Atheism, and the number of comments in each class is (C1:706 comments, C2:700 comments, C3:643 comments, and C4:716 comments).

6.2.5 Classifying extremist religious texts: Results and discussions

This work is a comparative study the application of eighteen (18) cases to classify the composite data set from five data sets. We train and test four classifiers on each case. We may characterize and rate results based on the accuracy and confusion matrix of the best classifier.

6.2.5.1 Classifying extremist religious texts: Results

The accuracy results of the four classifiers, support vector machines (SVM), decision tree (DT), random forest (RF), and Naive Bayes (NB), on the data sets in 18 Tests. The following table shows the results of the various tests.

Red	Maximum accuracy value of the test: The best result
Green	Minimum accuracy value of the test: The worst result

Table 6.23: Classification results.

Test	Class type & Nb comments	Class	Acc%	Max & Min accuracy in test
01	C1: includes Islamic comments, C2: the rest of the religious and non-religious comments. (C1:700, C2:700)	SVM	90.03	Maximum accuracy is 90.03% by SVM. Minimum accuracy is 88.25% by NB.
		DT	89.32	
		RF	89.67	
		NB	88.25	
02	C1: includes Christian comments, C2: includes the rest of the religious comments. (C1:643, C2:643)	SVM	86.04	Maximum accuracy is 86.04% by SVM and NB. Minimum accuracy is 84.49% by RF
		DT	85.65	
		RF	84.49	
		NB	86 .04	
03	C1: includes Christian comments, C2: includes the rest of the religious and non-religious comments. (C1:643, C2:643)	SVM	87.54	Maximum accuracy is 88.32% given by NB, Minimum accuracy is 82.87% given by DT
		DT	82.87	
		RF	84.04	
		NB	88.32	
04	C1: includes Islamic comments and C2: includes the rest of the religious comments. (C1:700, C2:700).	SVM	86.59	Maximum accuracy is 86.59% given by SVM. Minimum accuracy is 82.24% given by RF.
		DT	84.05	
		RF	82. 24	
		NB	85.86	
05	C1: includes Atheism comments, C2: includes the rest of the religious comments. (C1:706, C2:706)	SVM	75.97	Maximum accuracy is 78.09% given by NB and Minimum accuracy is 74.20% given by RF.
		DT	75.26	
		RF	74.20	
		NB	78.09	
06	C1: includes Atheism comments, C2: includes the rest of the religious and non-religious comments. (C1:706, C2:706)	SVM	68.79	Maximum accuracy is 68.43% given by NB and Minimum accuracy is 66.31% given by DT
		DT	66.31	

Continued on next page

Table 6.23 – continued from previous page

Test	Class type & Nb comments	Class	Acc%	Max Min acc in test
		RF	67.02	
		NB	68.43	
07	Two classes: C1: includes Jewish comments, C2: includes the rest of the religious and non-religious comments. (C1:716,C2:716)	SVM	69.47	Maximum accuracy is 69.47% given by SVM, Minimum accuracy is 60% given by NB
		DT	68.42	
		RF	67.01	
		NB	60	
08	C1: includes Jewish comments, C2: includes the rest of the religious comments. (C1:716, C2:716).	SVM	58.78	Maximum accuracy is 58.78% given by SVM. Minimum accuracy is 54.83% given by DT
		DT	54.83	
		RF	57.70	
		NB	58.06	
09	C1: includes Islamic comments, C2: includes the rest of the religious and non-religious.(C1:700, C2:2673)	SVM	93.33	Maximum accuracy is 93.77% given by RF, Minimum accuracy is 90.96% given by NB
		DT	92.88	
		RF	93.77	
		NB	90.96	
10	C1: includes Christian comments, C2: includes the rest of the religious and non-religious comments. (C1:643, C2:2730)	SVM	92.88	Maximum accuracy is 93.33% given by the RF. Minimum accuracy is 91.11% given by NB
		DT	92.44	
		RF	93.33	
		NB	91.11	
11	C1: includes Islamic comments, C2: includes the rest of the religious comments. (C1:700,C2:2065)	SVM	91.33	Maximum accuracy is 91.33% given by SVM, Minimum accuracy 89.71% given by DT
		DT	89.71	
		RF	90.79	
		NB	90.07	
12	C1: includes Christian comments , and C2: includes the rest of the religious comments. (C1:643, C2:2122)	SVM	81.94	Maximum accuracy is 81.94% given by SVM, Minimum accuracy is 74.18% given by DT.
		DT	74.18	

Continued on next page

Table 6.23 – continued from previous page

Test	Class type & Nb comments	Class	Acc%	Max Min acc in test
		RF	76.17	
		NB	78.70	
13	C1: includes Jewish comments, C2: includes the rest of the religious and non-religious comments. (C1:716, C2:2657)	SVM	80.14	Maximum accuracy is 80.14% given by SVM, Minimum accuracy is 66.96% given by NB
		DT	77.62	
		RF	78.66	
		NB	66.96	
14	C1 includes Atheism comments, C2: includes the rest of the religious and non-religious comments. (C1:706, C2:2667)	SVM	78.37	Maximum accuracy is 78.37% given by SVM. Minimum accuracy is 70.51% given by NB
		DT	75.70	
		RF	76.44	
		NB	70.51	
15	C1: includes Jewish comments, C2: includes the rest of the religious comments. (C1:716, C2:2049)	SVM	67.87	Maximum accuracy is 67.87% given by SVM. Minimum accuracy is 58.12% given by NB
		DT	60.64	
		RF	60.46	
		NB	58 .12	
16	C1: includes Atheism comments, C2: includes the rest of the religious comments. (C1:706, C2:2059)	SVM	64.98	Maximum accuracy is 64.98% given by SVM. Minimum accuracy is 55.95% given by NB
		DT	56.85	
		RF	58.48	
		NB	55.95	
17	5 classes (Islamic, Christian, Jewish, Atheism, neutral), number of comments (C1:706, C2:700, C3:643, C4:716, C5:608).	SVM	55.40	Maximum accuracy is 56.44% given by RF. Minimum accuracy is 46.81% given by NB
		DT	54.07	
		RF	56.44	
		NB	46.81	
18	4 classes: (Islamic, Christian, Jewish, and Atheism), number of comments (C1:706, C2:700, C3:643, C4:716)	SVM	53.79	Maximum accuracy is 53.79% given by SVM, Minimum accuracy is 46.57% given by RF
		DT	49.81	

Continued on next page

Table 6.23 – continued from previous page

Test	Class type & Nb comments	Class	Acc%	Max Min acc in test
		RF	46.57	
		NB	53.61	

6.2.5.2 Classifying extremist religious texts: Discussions

Through the results of the above table, the best result in all tests is (93.77%) by the RF in test (9):

- The number of comments in each of the first and second classes is unbalanced), this is the same for SVM and DT, their best results were with the ninth test (93.33%, 92.88%) respectively, but NB reached its maximum metrics (91.11%) in test (10).

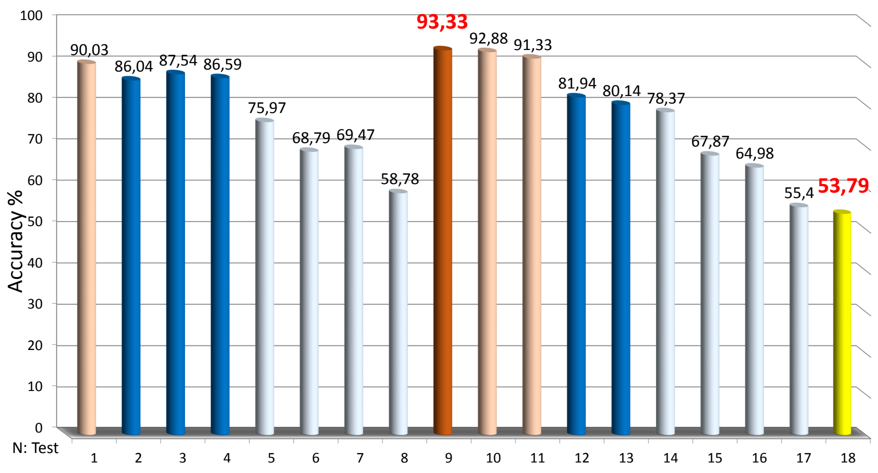

Figure 6.13: SVM classifier tests accuracy

- The dataset in this test is unbalanced; this is the same for SVM and DT, their best results were with the ninth test (93.33%, 92.88%) respectively, but NB reached its maximum accuracy (91.11%) in test (10).
- Minimum accuracy value of the test is (46.57%), is given by the RF with test 18, this is the same for SVM, DT, their worst results were with the test (53.79%), (49.81%) respectively Figure (6.14), but NB reached its minimum measurement (46.81%) in the test (17) Figure (6.15).

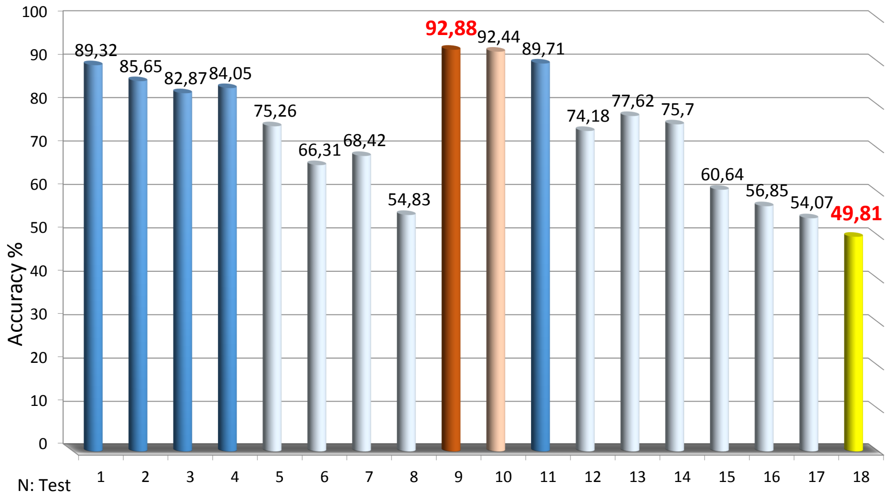

Figure 6.14: DT classifier tests accuracy

- In tests (17) and (18), we found that adding a class of neutral comments to the rest of the religious classes improved the results of classifiers except for NB, which decreased the results.
- Based on this, we can say that the size of the data has a greater impact on the improvement of the result than the number of classes.
- In tests (4) and (11), (8) and (15), we found that increasing the number of comments on the functionality of (test 4 and test 15) improved the results

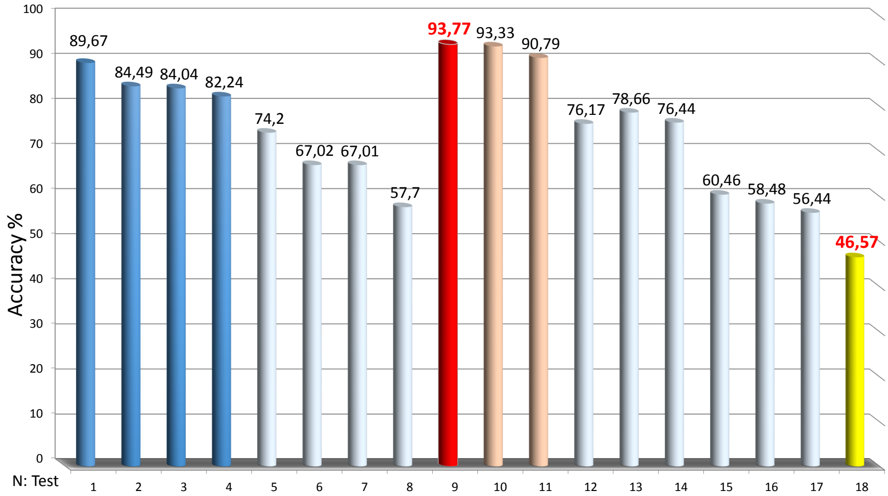

Figure 6.15: RF classifier tests accuracy

of all classifiers. In contrast to tests (2) and (12), (5) and (16), we found that the increase in the number of comments to the functionality of (test 2 and test 5) reduced the results of all the classifiers.

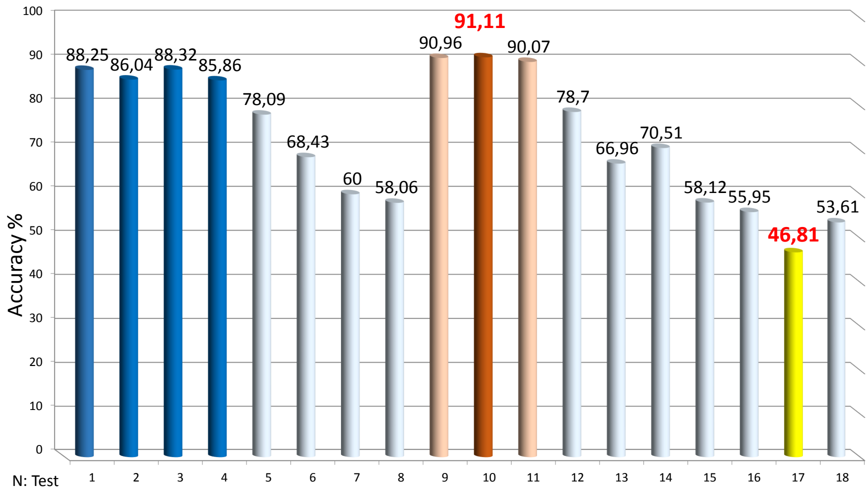

Figure 6.16: NB classifier tests accuracy

- In tests (1) and (9), (3) and (10), (7) and (13), (6) and (14), we found that

Increasing the number of comments to the functionality of (test 1, test 3, test 7, test 6) improved the results of all classifiers.

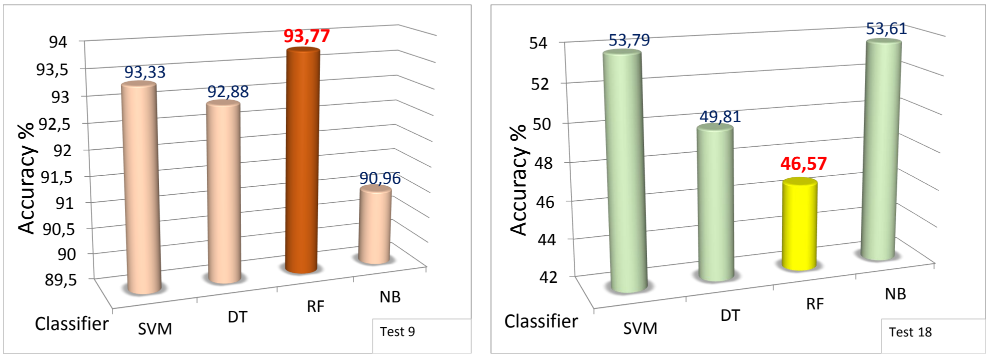

Figure 6.17: The best and the worst test of the proposed approach.

- Based on tests (4, 11, 2, 12, 8, 15, 5, and 16), we have observed that the balance and imbalance of the number of comments have the same effect. On the other hand, in tests (1, 9, 3, 10, 7, 13, 6, and 14), we noted that the balance and imbalance of the number of comments did not have the same effect.

From all these results for the different tests, it is clear that the final test (9) gave the best results, and we prefer this to determine and organize the classification of the data set (C1: includes Islamic comments, C2: includes the rest of the religious and non-religious comments). The basis for this is that the Islamic Commentaries have the most significant number of dictionaries of various characteristics (8 dictionaries).

We conclude that improving the result is not due to balance or the large dataset size but instead to the diversity of dictionaries of text features that restrict the text to its sentiment and linguistic framework.

6.2.6 Confusion matrix for Extremism detection and classification

From the confusion matrix in Figure (Ref fig:6.18), we notice that the predicted value of the first target variable (extreme text detection accuracy prediction values) is less than the actual value of the same target variable. However, the predicted value of the first target variable matches the actual value of the same target variable of the second and third levels, as seen in Figures (6.20 and 6.22).

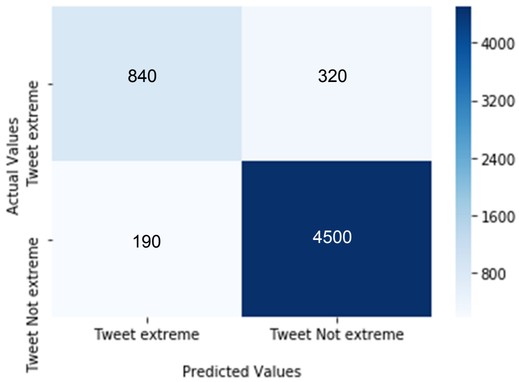

Figure 6.18: Confusion Matrix of Extremism class level L1 .

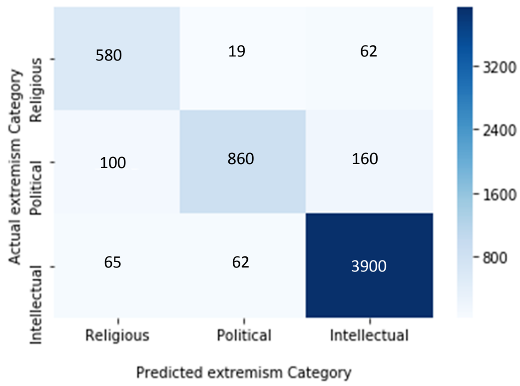

Figure 6.19: Confusion matrix to classify the nature of extremes level L1 .

The predicted value of the second target variable (Non-extreme text detection accuracy prediction values) matches the actual value of the same target variable. The classifier accuracy result is more or less perfect with the Levels (L1, L2) dataset (see Figures 6.18 and 6.20). The predicted value of the first target variable is ideal at Level L3 (Ref fig:6.22).

Figures (6.19, Ref fig:6.21 and 6.23) show the uncertainty output is fragile compared to the number of overlapping classes, but the confusion matrix for classifying multiple classes shows apparent confusion between the third, first, second and first classes. This is because the target variable's predicted value not matching the target variable's actual value is weak.

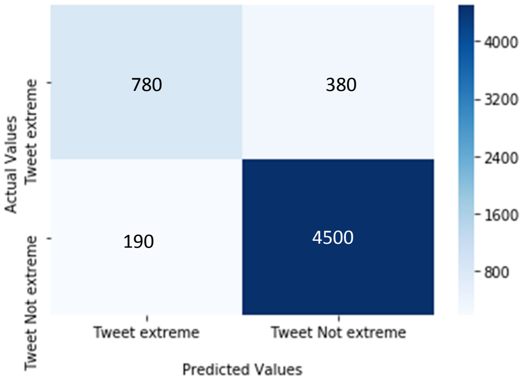

Figure 6.20: Confusion Matrix of Extremism class level L2 .

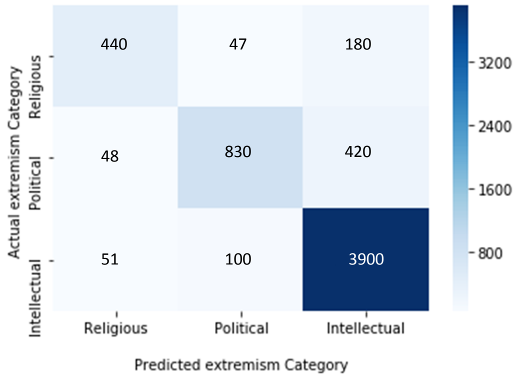

Figure 6.21: Confusion matrix to classify the nature of extremes level L2 .

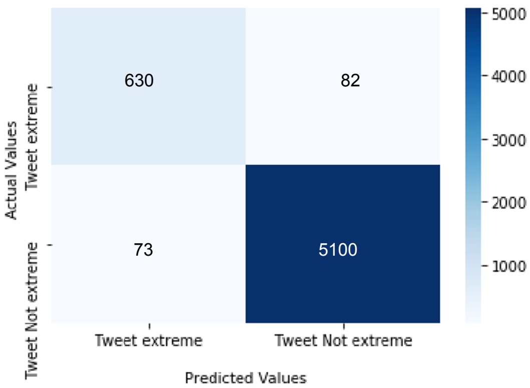

Figure 6.22: Confusion Matrix of Extremism class level L3.

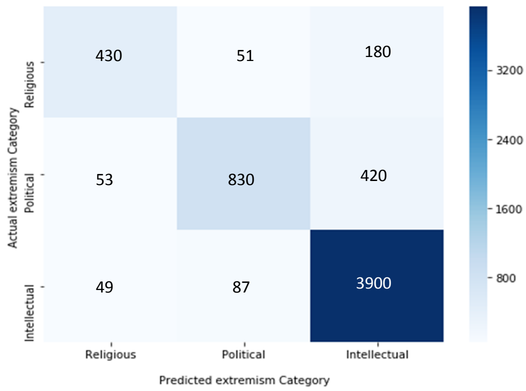

Figure 6.23: Confusion matrix on classifying the nature of extremes level L3.

The reason is that classifying tweets based on the type of tweet text is the basis of the problem because the tweets of the three categories mainly overlap in the first and second layers. This reduces certainty and increases confusion about the second-ranked result. The basis of extremist tweets is purely intellectual, but the similarity and convergence of some religious or political terms direct the tweet to other layers.

6.3 Comparison of literature related to our work

In Table (6.24), we compare the classification accuracy results in our work with those of the relevant literature.

Table 6.24: Comparison of extremism classification accuracy results for related literature.

Literature	Technique(s) NLP	Technique(s) ML	Accuracy
Our work	Models (1 & 2): TF-IDF with padding words dictionary and model (3): TF with feature dictionaries	Models (1 & 2): 16 machine learning classifiers and model (3): 4 machine learning	Acc-1= 97,00%, Acc-2= 93,70%, Acc-3= 93,77%
Matthew C. Benigni et al. [95].	Iterative Vertex Clustering and Classification (IVCC)	Common algorithms include K-means, Newman and Louvain Grouping	Acc: 96%
Mussiraliyeva Shynar et al. [97]	Word2Vec and TF-IDF	Gradient boosting and Random forest	Acc-1= 89%, Acc-2= 87%
Aditi Sarker et al. [80]	NRC and Word2Vec	Integrates sentiment-specific word embeddings and a weighted text feature model Vs. SSWE and WTFM1	Acc = 68,31 %
Sharif Waqas et al. [96]	N-gram , TF-IDF	SVM	Acc = 0,84%
Quanzhi Li et al. [142]	Bag of Word (BOW)	Ensemble of classifier	Acc = 82,2%
Omar Sharif et al. [143]	Bag of Word, TF-IDF and N-gram	LR, DT RF, MNB and SGD	Acc = 84,57%
Zia Ul Rehman et al. [144]	Word2Vec, TFIDF , FT, FP and FB	Random Forest classifiers (RFC)	Acc-1= 94%, Acc-2= 91%

We evaluated two extremism detection, and assessment models (5, 885 tweets, 20% of the dataset) used during the testing phase. The two models produced by our work can provide accurate predictions and good classification into different classes. Table 6.24 highlights the proposed natural language processing and classifier ML and demonstrates the accuracy of prediction and classification generated by NLP text feature extraction techniques and machine learning classifiers for our work and related literature. Other literature has relied on different intelligent approaches. We developed and refined an intelligent approach to natural language processing by utilizing text data processing techniques to make the dataset more visible and accurate in terms of focusing its features on extremes. By omitting the high-frequency padding words 6.12, which affect the accuracy results of the classifiers; By its effect on the quality and number of text features as shown by word clouds of Figure (6.8 , 6.9 and 6.10). It is eligible for use with intelligent natural language processing technologies. Across this data set, we use natural language processing techniques and machine learning methods to find the optimal classifier for detecting and categorizing extreme texts in a wide range of robust, high-resolution data.

According to the obtained results, the reached accuracy in this work competes with the case of the methods mentioned in Table (6.24). The main reason is that we used word clouds to show the most frequently occurring words in the dataset, and we cleansed the dataset of high-frequency padding words to make the dataset word features more focused on phrasing. Using Machine Learning classifiers (16 Machine Learning Classifiers) gives us an advantage in finding the best machine learning classifier using well-fixed metrics: accuracy, F1-score, MSE, and confusion matrix. The linear parameter SVC [137; 145] gives ideal results for the metrics in our work, the properties of which are manually manipulated.

This case makes our work complementary to the relevant literature to give it an idea of purging the dataset of padding words to get the best results in their studies.

6.4 Conclusion

This chapter presents an intelligent approach to three models of extremist text detection and classification (religious, political, and intellectual) and religious extremist text classification. To create the first and second models, data collection, dataset preprocessing, text feature extraction (TF-IDF), and a set of 16 machine learning classifiers were used to improve the accuracy of extreme text detection and classification. The features of the extremist religious text were also extracted and collected in dictionaries with (the TF) technique of NLP and a group of four Machine Learning classifiers. The application of main contributions of the proposed approach was:

- First: Preprocessing a raw dataset containing (29,423) tweets to discover the extremist text and classify its types.
- Second: Work on extracting feature dictionaries for a dataset containing comments (3373) to classify religious communities. To classify the extremist religious text.

To detect text extremism and classify its types, an intelligent prediction classifier was developed using data set prepossessing 16 machine learning classifiers at three levels. The SVC linear classifier was given an intelligent extreme text classification model with high accuracy (97.7%). After applying the bagging word cleanup function, we found that the SVC linear classifier given an intelligent extreme text detection model achieved high accuracy (97%).

An intelligent taxonomic classifier was developed to classify the religious community using data collection, inferring a data set, and extracting its features in dictionaries using (TF) technique, training, and testing machine learning classifiers. We found that the Random Forest (RF) classifier achieved significant accuracy (93.77%).

Chapter 7

General conclusion

7.1 Conclusions

Recently, the social web has developed rapidly and effectively. The number of its users exceeded (3 billion) out of the (4.54 billion) users on the Internet [146]. The analysis and processing of social network data have become imperative for analyzing the sentiment and desires of its users. The text had the largest share of studies, research, and analysis because it was more widespread and influential. The impact and effect of text information on social networks spread incredibly abusive and negative.

The spread of extremist ideas and ideologies, especially hate speech and extremist texts with fanatical and extremist religious, political, or intellectual ideas and opinions on social media, and the violation of all laws, legislation, and customs of people led to the recruitment of individuals and groups. To counter this dark side of social networks, this thesis presented an intelligent approach to detecting extremist text, classifying its type, and classifying extremist religious text through extremist texts on social networks. Moreover, save a considerable amount of data, especially extreme text data.

The thesis included six parts. The first part is an introduction to the scope of the research, its main problems, challenges, and the goal to be reached. Presenting our contribution to the theoretical and practical aspects and the techniques and approaches used in them. The second part, State of the Art: chapters two and three present definitions of extremism, its concept, and its relationship to social networks. We also referred to text mining and analysis through automated natural language processing. The fourth part contains the fourth chapter. We presented the structure of the proposed curriculum and its basic levels in general. The results and their discussion in Chapter Six were an application of the three proposals to the proposed structure, and they are fixed values and relate to the proposed data set and the proposed approach.

Contributions

The first applied contribution to the first theoretical contribution was presented in the form of an application to a set of (16)sixteen machine learning classifiers at three levels for the proposed dataset to detect extremist text in the first stage and to classify the extremist text in the second stage:

- **The first level:** Is the application of the raw dataset as it is without change.
- **The second level:** Is the application to the prepossessing data set by applying a set of functions based on the prepossessing of the raw textual data set related to extremism. These functions make the data set more transparent and accurate.
- **The third level:** We added to the prepossessing of the dataset a new function to remove high-frequency padding words based on two main elements in their input:
    - The raw dataset and its main source is the social network.

– High-Frequency Padding Word Dictionary: It counts the frequency of all the raw dataset words and takes from them the high-frequency words (meaningless, not basic features, incomprehensible).

After implementing this function, dataset terminology became more precise regarding focusing its features on extremism.

The second applied contribution to the second theoretical contribution was presented in the form of an application of a set of (4)four machine learning classifiers on (18) eighteen tests of the extremist text dataset of different types of religions (Christian, Jewish, Muslim and atheist with the addition of plain text). Dictionaries were used to distinguish features of each religion.

The second applied contribution was made by applying an intelligent approach to natural language processing using textual data processing techniques by processing the first datasets (Twitter) in two stages: Stage 1: We use TF-IDF (NLP technology) to extract text features and convert them into digital vectors. Stage 2: We use TF (NLP technology) with feature dictionaries to output text features and convert them into digital vectors.

The third applied contribution was made in two phases:

- **The first phase:** We implemented and evaluated sixteen intelligent classified algorithms for machine learning on the extreme text dataset. We obtained very high accuracy (97%) using the SVC linear classifier to detect extremes and very high accuracy (93.6%) using the SGD classifier to classify extremes.

- **The second phase:** We implemented and evaluated four intelligent classifier algorithms for machine learning on the extreme text dataset of instruments. In eighteen tests with different dictionaries during the teaching and testing of the proposed machine learning (SVM, DT, RF, NB). These tests gave acceptable results, the best for the SVM Accurate Linear Classifier (93.77%).

The proposed approaches have proven to be effective in improving the results of metrics for evaluating machine learning classifiers. The first classifier for extreme text detection showed improvement from level L1 to level L3 with percentages as follows: training accuracy of (0.3%), testing accuracy of (10.9%), F1_score of (11.3%), and MSE of (−18.6%). The second classifier for extreme text type classification showed improvement from level L1 to level L3 with percentages as follows: (training accuracy of (0.4%), testing accuracy of (2.6%), F1_score 4.8% and MSE of (−8.8%)). The third classifier for the classification of the extremist religious text achieved the best accuracy for the ninth test out of a total of eighteen tests with a percentage of (93.77%).

The relationship between the religious text and the text of extremism is optional. There are types of non-religious extremism. Other reasons are intellectual, political, social, and economic. The concept of religion must be separated from extremism, as the various religious texts spread in social networks contain a lot of goodness and mercy, and few of them are extremists. Caution this little significantly negatively impacts the owners of fanatical and extremist ideas. In addition, many forms of extremism are not related to religion but to other causes. Religion must remain a source of mercy and love in this world, and we must separate it from extremism for the sake of humanity in our world.

7.2 Future works

In the future, we will work on other axes in the context. We will work on other languages to extract the features of religious, politique, intellectual classifications, and extremism features in their texts using multilingual dynamic social groups analysis techniques[147].

We will also work on using other criteria for detection and classification, such as educational levels and living levels and their relationship with extremism.

This approach can also be exploited in sensitive fields such as psychiatry (Ex: Alzheimer's) [148; 149], which relies on dialogues and audio recordings that can be converted into texts and then studied and analyzed.

Bibliographies

[1] H. Kaur, V. Mangat, and Nidhi, "A survey of sentiment analysis techniques," in *2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, pp. 921–925, 2017. vii, 46

[2] C. Zucco, B. Calabrese, G. Agapito, P. H. Guzzi, and M. Cannataro, "Sentiment analysis for mining texts and social networks data: Methods and tools," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 10, no. 1, p. e1333, 2020. 1

[3] D. Crystal *et al.*, *English as a global language.* Cambridge university press, 2003. 1

[4] H. Allcott, L. Braghieri, S. Eichmeyer, and M. Gentzkow, "The welfare effects of social media," *American Economic Review*, vol. 110, no. 3, pp. 629–76, 2020. 1

[5] L. Milroy and C. Llamas, "Social networks," *The handbook of language variation and change*, pp. 407–427, 2013. 1

[6] B. C. Kelly and M. Vuolo, "Social network ties to nightlife and healthcare professionals and prescription drug misuse among young adults," *International Journal of Drug Policy*, vol. 66, pp. 48–56, 2019. 2

[7] K. Clark, J. B. Fletcher, I. W. Holloway, and C. J. Reback, "Structural in-

equities and social networks impact hormone use and misuse among transgender women in los angeles county," *Archives of sexual behavior*, vol. 47, no. 4, pp. 953–962, 2018. 2

[8] I. Jugl, F. Lösel, D. Bender, and S. King, "Psychosocial prevention programs against radicalization and extremism: a meta-analysis of outcome evaluations," *European journal of psychology applied to legal context*, vol. 13, no. 1, pp. 37–46, 2020. 3

[9] X. Xie, Y. Fu, H. Jin, Y. Zhao, and W. Cao, "A novel text mining approach for scholar information extraction from web content in chinese," *Future Generation Computer Systems*, vol. 111, pp. 859–872, 2020. 3, 30

[10] A. Berhoum, M. C. E. Meftah, A. Laouid, and M. Hammoudeh, "An intelligent approach based on cleaning up of inutile contents for extremism detection and classification in social networks," *ACM Trans. Asian Low-Resour. Lang. Inf. Process.*, jan 2023. Just Accepted. 9

[11] A. Berhoum, M. C. E. Meftah, A. Laouid, and M. Hammoudeh, "Machine learning to classify religious communities and detect extremism on social networks: Ml to crcs and de through text tweets on sns," *International Journal of Organizational and Collective Intelligence (IJOCI)*, vol. 12, no. 1, pp. 1–19, 2022. 9

[12] V. Amelkin and A. K. Singh, "Fighting opinion control in social networks via link recommendation," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 677–685, 2019. 12

[13] a. M. Haider Khudair Murad Lafta, "Historical roots of extremism in con-

temporary islamic terrorist organizations," *Al-Bahith Journal*, vol. 36, no. 1, 2020. 12

[14] V. Badaan and J. T. Jost, "Conceptual, empirical, and practical problems with the claim that intolerance, prejudice, and discrimination are equivalent on the political left and right," *Current Opinion in Behavioral Sciences*, vol. 34, pp. 229–238, 2020. 12

[15] G. C. J. Ivan, "From ideology to hate speech and the problem of euro-american white supremacist extremism," *Bulletin of the Perm University. Series: Political Science*, pp. 5–13, 2020. 12

[16] S. Wibisono, W. R. Louis, and J. Jetten, "A multidimensional analysis of religious extremism," *Frontiers in psychology*, vol. 10, p. 2560, 2019. 12

[17] A. Sotlar, "Some problems with a definition and perception of extremism within a society," *Policing in central and Eastern Europe: Dilemmas of contemporary criminal justice*, pp. 703–707, 2004. 12

[18] S. Aldera, A. Emam, M. Al-Qurishi, M. Alrubaian, and A. Alothaim, "Online extremism detection in textual content: A systematic literature review," *IEEE Access*, vol. 9, pp. 42384–42396, 2021. 12

[19] S. Trip, C. H. Bora, M. Marian, A. Halmajan, and M. I. Drugas, "Psychological mechanisms involved in radicalization and extremism. a rational emotive behavioral conceptualization," *Frontiers in psychology*, vol. 10, p. 437, 2019. 12

[20] S. A. Hamdi, "Mining ideological discourse on twitter: The case of extremism in arabic," *Discourse & Communication*, vol. 16, no. 1, pp. 76–92, 2022. 12

[21] C. Allen, A. Isakjee, and Ö. Ögtem-Young, "Counter-extremism, prevent and the extreme right wing: Lessons learned and future challenges," *LIAS Working Paper Series*, vol. 2, 2019. 12

[22] K. Sharma, "What causes extremist attitudes among sunni and shia youth? evidence from northern india," in *Evidence from Northern India*, Program on Extremism, 2016. 12, 13

[23] J. M. Lutz, "Extremism: by jm berger, cambridge, ma, mit press, 2018, 201+ x pp., $15.95/£ 11.95," *The European Legacy*, vol. 26, no. 3-4, 2020. 13

[24] C. Bott, W. J. Castan, R. Dickens, T. Rowley, E. Smith, R. Lark, and G. Thompson, "Recruitment and radicalization of school aged youth by international terrorist groups," *Final Report. Online verfügbar: https://pdfs. semanticscholar. org/c664/9103b3d86b13fa58b2c1a2942b5b7ada91d7. pdf.(Stand: 07.12. 2017)*, 2009. 13

[25] A. Van de Weert and Q. A. Eijkman, "Subjectivity in detection of radicalisation and violent extremism: a youth worker's perspective," *Behavioral Sciences of Terrorism and Political Aggression*, vol. 11, no. 3, pp. 191–214, 2019. 13

[26] A. P. Schmid, "Radicalisation, de-radicalisation, counter-radicalisation: A conceptual discussion and literature review," *ICCT Research Paper*, vol. 97, no. 1, p. 22, 2013. 13

[27] B. Schuurman and M. Taylor, "Reconsidering radicalization: Fanaticism and the link between ideas and violence," *Perspectives on Terrorism*, vol. 12, no. 1, pp. 3–22, 2018. 14

[28] A. Bötticher, "Towards academic consensus definitions of radicalism and extremism," *Perspectives on terrorism*, vol. 11, no. 4, pp. 73–77, 2017. 14

[29] M. Mondal, L. A. Silva, and F. Benevenuto, "A measurement study of hate speech in social media," in *Proceedings of the 28th ACM conference on hypertext and social media*, pp. 85–94, 2017. 14

[30] P. Nilan, "Online discourse and social media," in *Young People and the Far Right*, pp. 29–56, Springer, 2021. 14

[31] M. Khosravi Nik and M. W. Amer, "Social media and terrorism discourse: the islamic state's (is) social media discursive content and practices," *Critical Discourse Studies*, vol. 19, no. 2, pp. 124–143, 2022. 14

[32] V. X. Nguyen, G. Xiao, J. Zhou, G. Li, and B. Li, "Bias in social interactions and emergence of extremism in complex social networks," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 30, no. 10, p. 103110, 2020. 14

[33] E. Anwas, Y. Sugiarti, A. Permatasari, J. Warsihna, Z. Anas, L. Alhapip, H. Siswanto, and R. Rivalina, "Social media usage for enhancing english language skill," 2020. 15

[34] M. Muftah, "Impact of social media on learning english language during the covid-19 pandemic," *PSU Research Review*, 2022. 15

[35] R. Irfan, C. K. King, D. Grages, S. Ewen, S. U. Khan, S. A. Madani, J. Kolodziej, L. Wang, D. Chen, A. Rayes, *et al.*, "A survey on text mining in social networks," *The Knowledge Engineering Review*, vol. 30, no. 2, pp. 157–170, 2015. 15, 16

[36] P. J. Stone, "Thematic text analysis: New agendas for analyzing text content," *Text analysis for the social sciences*, pp. 35–54, 2020. 15

[37] A. Hotho, A. Nürnberger, and G. Paaß, "A brief survey of text mining.," in *Ldv Forum*, pp. 19–62, Citeseer, 2005. 16

[38] L. Zhao, F. Chen, J. Dai, T. Hua, C.-T. Lu, and N. Ramakrishnan, "Unsupervised spatial event detection in targeted domains with applications to civil unrest modeling," *PloS one*, vol. 9, no. 10, p. e110206, 2014. 16, 39

[39] N. Naveed, T. Gottron, J. Kunegis, and A. C. Alhadi, "Bad news travel fast: A content-based analysis of interestingness on twitter," in *Proceedings of the 3rd international web science conference*, pp. 1–7, 2011. 16, 39

[40] S. B. B. Priyadarshini, A. B. Bagjadab, and B. K. Mishra, "A brief overview of natural language processing and artificial intelligence," *Natural Language Processing in Artificial Intelligence*, pp. 211–224, 2020. 16

[41] S. Kannan, V. Gurusamy, S. Vijayarani, J. Ilamathi, M. Nithya, S. Kannan, and V. Gurusamy, "Preprocessing techniques for text mining," *International Journal of Computer Science & Communication Networks*, vol. 5, no. 1, pp. 7–16, 2014. 17

[42] S. Vijayarani, M. J. Ilamathi, M. Nithya, *et al.*, "Preprocessing techniques for text mining-an overview," *International Journal of Computer Science & Communication Networks*, vol. 5, no. 1, pp. 7–16, 2015. 17

[43] X. Zhang, J. Cui, W. Wang, and C. Lin, "A study for texture feature extraction of high-resolution satellite images based on a direction measure and gray level co-occurrence matrix fusion algorithm," *Sensors*, vol. 17, no. 7, p. 1474, 2017. 18

[44] U. Gadiraju, R. Yu, S. Dietze, and P. Holtz, "Analyzing knowledge gain of users in informational search sessions on the web," in *Proceedings of the 2018 Conference on Human Information Interaction & Retrieval*, pp. 2–11, 2018. 18

[45] G. Sidorov, F. Velasquez, E. Stamatatos, A. Gelbukh, and L. Chanona-Hernández, "Syntactic dependency-based n-grams as classification features," in *Mexican International Conference on Artificial Intelligence*, pp. 1–11, Springer, 2012. 18

[46] L. G. Smith, L. Wakeford, T. F. Cribbin, J. Barnett, and W. K. Hou, "Detecting psychological change through mobilizing interactions and changes in extremist linguistic style," *Computers in Human Behavior*, vol. 108, p. 106298, 2020. 18, 47

[47] R.-C. Chen, J.-Y. Liang, and R.-H. Pan, "Using recursive art network to construction domain ontology based on term frequency and inverse document frequency," *Expert systems with Applications*, vol. 34, no. 1, pp. 488–501, 2008. 18

[48] V. Y. Tsvetkov, "Dichotomous systemic analysis," *Life Science Journal*, vol. 11, no. 6, pp. 586–590, 2014. 18

[49] S. Le Cessie and J. C. Van Houwelingen, "Ridge estimators in logistic regression," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 41, no. 1, pp. 191–201, 1992. 18

[50] X. Ma, P. Zhang, S. Zhang, N. Duan, Y. Hou, M. Zhou, and D. Song, "A tensorized transformer for language modeling," *Advances in neural information processing systems*, vol. 32, 2019. 18

[51] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, and S. Zagoruyko, "End-to-end object detection with transformers," in *European conference on computer vision*, pp. 213–229, Springer, 2020. 18

[52] O. Levy and Y. Goldberg, "Dependency-based word embeddings," in *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pp. 302–308, 2014. 18

[53] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *arXiv preprint arXiv:1301.3781*, 2013. 19

[54] P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov, "Enriching word vectors with subword information," *Transactions of the association for computational linguistics*, vol. 5, pp. 135–146, 2017. 19

[55] J. Pennington, R. Socher, and C. D. Manning, "Glove: Global vectors for word representation," in *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pp. 1532–1543, 2014. 19

[56] K. Krippendorff, *Content analysis: An introduction to its methodology.* Sage publications, 2018. 19

[57] A. M. Turing, "Computing machinery and intelligence," in *Parsing the turing test*, pp. 23–65, Springer, 2009. 21

[58] R. Jackendoff, "Parallel constraint-based generative theories of language," *Trends in Cognitive Sciences*, vol. 3, no. 10, pp. 393–400, 1999. 21

[59] M. P. Deisenroth, A. A. Faisal, and C. S. Ong, *Mathematics for machine learning.* Cambridge University Press, 2020. 22

[60] E. Kauffmann, J. Peral, D. Gil, A. Ferrández, R. Sellers, and H. Mora, "Managing marketing decision-making with sentiment analysis: An evaluation of the main product features using text data mining," *Sustainability*, vol. 11, no. 15, p. 4235, 2019. 22

[61] Y. Zhang and Z. Teng, *Natural language processing: a machine learning perspective*. Cambridge University Press, 2021. 22

[62] A. J. Trappey, C. V. Trappey, J.-L. Wu, and J. W. Wang, "Intelligent compilation of patent summaries using machine learning and natural language processing techniques," *Advanced Engineering Informatics*, vol. 43, p. 101027, 2020. 22

[63] M. Razno, "Machine learning text classification model with nlp approach," in *COLINS 2019. Volume II: Workshop*, 2019. 23

[64] S. Jaeger, S. Fulle, and S. Turk, "Mol2vec: unsupervised machine learning approach with chemical intuition," *Journal of chemical information and modeling*, vol. 58, no. 1, pp. 27–35, 2018. 23

[65] Z. Alizadeh-Sani, P. P. Martínez, G. H. González, A. González-Briones, P. Chamoso, and J. M. Corchado, "A hybrid supervised/unsupervised machine learning approach to classify web services," in *Practical Applications of Agents and Multi-Agent Systems*, pp. 93–103, Springer, 2021. 23

[66] T. Jiang, J. L. Gradus, and A. J. Rosellini, "Supervised machine learning: a brief primer," *Behavior Therapy*, vol. 51, no. 5, pp. 675–687, 2020. 24

[67] T. Pranckevičius and V. Marcinkevičius, "Comparison of naive bayes, random forest, decision tree, support vector machines, and logistic regression classifiers for text reviews classification," *Baltic Journal of Modern Computing*, vol. 5, no. 2, p. 221, 2017. 24

[68] J. Yousif and M. Al-Risi, "Part of speech tagger for arabic text based support vector machines: A review," *ICTACT Journal on Soft Computing: DOI*, vol. 10, 2019. 24

[69] S. M. Stigler, "Who discovered bayes's theorem?," *The American Statistician*, vol. 37, no. 4a, pp. 290–296, 1983. 25

[70] F. I. Adiba, T. Islam, M. S. Kaiser, M. Mahmud, and M. A. Rahman, "Effect of corpora on classification of fake news using naive bayes classifier," *International Journal of Automation, Artificial Intelligence and Machine Learning*, vol. 1, no. 1, pp. 80–92, 2020. 25

[71] S. Kunal, A. Saha, A. Varma, and V. Tiwari, "Textual dissection of live twitter reviews using naive bayes," *Procedia computer science*, vol. 132, pp. 307–313, 2018. 25

[72] T. Boros, S. D. Dumitrescu, and S. Pipa, "Fast and accurate decision trees for natural language processing tasks.," in *RANLP*, pp. 103–110, 2017. 25

[73] H. H. Patel and P. Prajapati, "Study and analysis of decision tree based classification algorithms," *International Journal of Computer Sciences and Engineering*, vol. 6, no. 10, pp. 74–78, 2018. 26

[74] J. Antony Vijay, H. Anwar Basha, and J. Arun Nehru, "A dynamic approach for detecting the fake news using random forest classifier and nlp," in *Computational methods and data engineering*, pp. 331–341, Springer, 2021. 26

[75] S. Agarwal and A. Sureka, "Applying social media intelligence for predicting and identifying on-line radicalization and civil unrest oriented threats," *arXiv preprint arXiv:1511.06858*, 2015. 28

[76] S. Liu, L. Zhang, and Z. Yan, "Predict pairwise trust based on machine learning in online social networks: A survey," *IEEE Access*, vol. 6, pp. 51297–51318, 2018. 30

[77] T. Taira, "The category of "religion" in public classification: Charity registration of the druid network (with suzanne owen)," in *Taking 'Religion'Seriously: Essays on the Discursive Study of Religion*, pp. 92–114, Brill, 2022. 30

[78] E. Ferrara, W.-Q. Wang, O. Varol, A. Flammini, and A. Galstyan, "Predicting online extremism, content adopters, and interaction reciprocity," in *International conference on social informatics*, pp. 22–39, Springer, 2016. 30

[79] S. A. Azizan and I. A. Aziz, "Terrorism detection based on sentiment analysis using machine learning," *Journal of Engineering and Applied Sciences*, vol. 12, no. 3, pp. 691–698, 2017. 30

[80] A. Sarker, P. Chakraborty, S. S. Sha, M. Khatun, M. R. Hasan, and K. Banerjee, "Improvised technique for analyzing data and detecting terrorist attack using machine learning approach based on twitter data," *Journal of Computer and Communications*, vol. 8, no. 7, pp. 50–62, 2020. 31, 93

[81] V. Chang, Y. Mou, Q. A. Xu, H. Kaur, and B. S. Liu, "Homicide network detection based on social network analysis.," in *IoTBDS*, pp. 329–337, 2021. 32

[82] M. L. Jamil, S. Pais, and J. Cordeiro, "Detection of dangerous events on social media: A perspective review," *arXiv preprint arXiv:2204.01351*, 2022. 32

[83] U. Bhattacharjee, P. Srijith, and M. S. Desarkar, "Term specific tf-idf boosting for detection of rumours in social networks," in *2019 11th International Conference on Communication Systems & Networks (COMSNETS)*, pp. 726–731, IEEE, 2019. 32

[84] M. M. Islam, M. A. Uddin, L. Islam, A. Akter, S. Sharmin, and U. K. Acharjee, "Cyberbullying detection on social networks using machine learning approaches," in *2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE)*, pp. 1–6, IEEE, 2020. 32

[85] W. Han and V. Mehta, "Fake news detection in social networks using machine learning and deep learning: Performance evaluation," in *2019 IEEE International Conference on Industrial Internet (ICII)*, pp. 375–380, IEEE, 2019. 33

[86] Á. de Pablo, Ó. Araque, and C. A. Iglesias, "Radical text detection based on stylometry," in *ICISSP*, 2020. 33

[87] L. Macnair and R. Frank, "Changes and stabilities in the language of islamic state magazines: A sentiment analysis," *Dynamics of Asymmetric Conflict*, vol. 11, no. 2, pp. 109–120, 2018. 33

[88] O. Klein and J. Muis, "Online discontent: Comparing western european far-right groups on facebook," *European societies*, vol. 21, no. 4, pp. 540–562, 2019. 33

[89] R. Scrivens, G. Davies, and R. Frank, "Measuring the evolution of radical right-wing posting behaviors online," *Deviant Behavior*, vol. 41, no. 2, pp. 216–232, 2020. 34

[90] L. Dillon, L. S. Neo, and J. D. Freilich, "A comparison of isis foreign fighters and supporters social media posts: An exploratory mixed-method content

analysis," *Behavioral sciences of terrorism and political aggression*, vol. 12, no. 4, pp. 268–291, 2020. 34

[91] A. I. Abd-Elaal, A. Z. Badr, and H. M. Mahdi, "Detecting violent radical accounts on twitter," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 8, 2020. 34, 35

[92] I. A. Farha and W. Magdy, "Mazajak: An online arabic sentiment analyser," in *Proceedings of the Fourth Arabic Natural Language Processing Workshop*, pp. 192–198, 2019. 35

[93] M. Ashcroft, A. Fisher, L. Kaati, E. Omer, and N. Prucha, "Detecting jihadist messages on twitter," in *2015 European intelligence and security informatics conference*, pp. 161–164, IEEE, 2015. 35

[94] H. Saif, T. Dickinson, L. Kastler, M. Fernandez, and H. Alani, "A semantic graph-based approach for radicalisation detection on social media," in *European semantic web conference*, pp. 571–587, Springer, 2017. 35, 36

[95] M. C. Benigni, K. Joseph, and K. M. Carley, "Online extremism and the communities that sustain it: Detecting the isis supporting community on twitter," *PloS one*, vol. 12, no. 12, p. e0181405, 2017. 35, 36, 93

[96] W. Sharif, S. Mumtaz, Z. Shafiq, O. Riaz, T. Ali, M. Husnain, and G. S. Choi, "An empirical approach for extreme behavior identification through tweets using machine learning," *Applied Sciences*, vol. 9, no. 18, p. 3723, 2019. 35, 36, 93

[97] S. Mussiraliyeva, M. Bolatbek, B. Omarov, Z. Medetbek, G. Baispay, and R. Ospanov, "On detecting online radicalization and extremism using natural language processing," in *2020 21st International Arab Conference on Information Technology (ACIT)*, pp. 1–5, IEEE, 2020. 35, 37, 93

[98] J. Eisenstein, *Introduction to natural language processing.* MIT press, 2019. 38

[99] R. Compton, C. Lee, J. Xu, L. Artieda-Moncada, T.-C. Lu, L. D. Silva, and M. Macy, "Using publicly visible social media to build detailed forecasts of civil unrest," *Security informatics*, vol. 3, no. 1, pp. 1–10, 2014. 41

[100] A. Twitter, "Twitter developer," *URL: https://developer. twitter. com/en/docs/twitter-api [accessed 2020-10-02]*, 2020. 41, 42

[101] M. N. Moghadasi, Z. Safari, and Y. Zhuang, "A sentimental and semantical analysis on facebook comments to detect latent patterns," in *2020 IEEE International Conference on Big Data (Big Data)*, pp. 4665–4671, IEEE, 2020. 41

[102] T. C. Bailey, A. C. Gatrell, *et al.*, *Interactive spatial data analysis.* Longman Scientific & Technical Essex, 1995. 41

[103] C. Zhang, J. Sun, X. Zhu, and Y. Fang, "Privacy and security for online social networks: challenges and opportunities," *IEEE network*, vol. 24, no. 4, pp. 13–18, 2010. 41

[104] N. Kewsuwun and S. Kajornkasirat, "A sentiment analysis model of agritech startup on facebook comments using naive bayes classifier.," *International Journal of Electrical & Computer Engineering (2088-8708)*, vol. 12, no. 3, 2022. 41

[105] J. M. Banda, R. Tekumalla, G. Wang, J. Yu, T. Liu, Y. Ding, E. Artemova, E. Tutubalina, and G. Chowell, "A large-scale covid-19 twitter chatter dataset for open scientific research—an international collaboration," *Epidemiologia*, vol. 2, no. 3, pp. 315–324, 2021. 42

[106] S. H. Mahmud, M. A. Hossin, H. Jahan, S. R. H. Noori, and T. Bhuiyan, "Csv-annotate: Generate annotated tables from csv file," in *2018 International Conference on Artificial Intelligence and Big Data (ICAIBD)*, pp. 71–75, IEEE, 2018. 42

[107] A. Nayak, B. Božić, and L. Longo, "Data quality assessment of comma separated values using linked data approach," in *International Conference on Business Information Systems*, pp. 240–250, Springer, 2022. 42

[108] O. Benjelloun, S. Chen, and N. Noy, "Google dataset search by the numbers," in *International Semantic Web Conference*, pp. 667–682, Springer, 2020. 42

[109] D. V. Sikuler, "Resources providing data for machine learning and ai validation," *Information and mathematical technologies in science and management*, no. 2 (22), pp. 39–52, 2021. 42

[110] C. S. Bojer and J. P. Meldgaard, "Kaggle forecasting competitions: An overlooked learning opportunity," *International Journal of Forecasting*, vol. 37, no. 2, pp. 587–603, 2021. 42

[111] D. E. White, N. D. Oelke, and S. Friesen, "Management of a large qualitative data set: Establishing trustworthiness of the data," *International journal of qualitative methods*, vol. 11, no. 3, pp. 244–258, 2012. 42

[112] J. Faigl, "Data collection path planning with spatially correlated measurements using growing self-organizing array," *Applied Soft Computing*, vol. 75, pp. 130–147, 2019. 42

[113] R. Ahuja, A. Chug, S. Kohli, S. Gupta, and P. Ahuja, "The impact of features extraction on the sentiment analysis," *Procedia Computer Science*, vol. 152, pp. 341–348, 2019. 44

[114] P. Tyagi and R. Tripathi, "A review towards the sentiment analysis techniques for the analysis of twitter data," in *Proceedings of 2nd international conference on advanced computing and software engineering (ICACSE)*, 2019. 44

[115] D. Ramachandran and R. Parvathi, "Analysis of twitter specific preprocessing technique for tweets," *Procedia Computer Science*, vol. 165, pp. 245–251, 2019. 45, 60

[116] A. Van Looy, "Sentiment analysis and opinion mining (business intelligence 1)," in *Social Media Management*, pp. 147–163, Springer, 2022. 45

[117] S. Qaiser and R. Ali, "Text mining: use of tf-idf to examine the relevance of words to documents," *International Journal of Computer Applications*, vol. 181, no. 1, pp. 25–29, 2018. 45, 46

[118] J. Ramos *et al.*, "Using tf-idf to determine word relevance in document queries," in *Proceedings of the first instructional conference on machine learning*, pp. 29–48, Citeseer, 2003. 46

[119] A. K. Singh and M. Shashi, "Vectorization of text documents for identifying unifiable news articles," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 7, 2019. 47

[120] N. R. Bhowmik, M. Arifuzzaman, M. R. H. Mondal, and M. Islam, "Bangla text sentiment analysis using supervised machine learning with extended lexicon dictionary," *Natural Language Processing Research*, vol. 1, no. 3-4, pp. 34–45, 2021. 47

[121] D. Deepa, A. Tamilarasi, *et al.*, "Sentiment analysis using feature extraction and dictionary-based approaches," in *2019 Third International confer-*

ence on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC), pp. 786–790, IEEE, 2019. 47

[122] O. I. per a la Normalització, *Accuracy (trueness and Precision) of Measurement Methods and Results*. International Organization for Standardization, 1994. 49

[123] D. M. Powers, "Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation," *arXiv preprint arXiv:2010.16061*, 2020. 49

[124] Z. Wang and A. C. Bovik, "Mean squared error: Love it or leave it? a new look at signal fidelity measures," *IEEE signal processing magazine*, vol. 26, no. 1, pp. 98–117, 2009. 49

[125] E. L. Lehmann and G. Casella, *Theory of point estimation*. Springer Science & Business Media, 2006. 49

[126] S. Visa, B. Ramsay, A. L. Ralescu, and E. Van Der Knaap, "Confusion matrix-based feature selection.," *MAICS*, vol. 710, no. 1, pp. 120–127, 2011. 50

[127] M. Heydarian, T. E. Doyle, and R. Samavi, "Mlcm: multi-label confusion matrix," *IEEE Access*, vol. 10, pp. 19083–19095, 2022. 50

[128] G. T. Reddy, S. Bhattacharya, S. S. Ramakrishnan, C. L. Chowdhary, S. Hakak, R. Kaluri, and M. P. K. Reddy, "An ensemble based machine learning model for diabetic retinopathy classification," in *2020 international conference on emerging trends in information technology and engineering (ic-ETITE)*, pp. 1–6, IEEE, 2020. 51

[129] A. Gholamy, V. Kreinovich, and O. Kosheleva, "Why 70/30 or 80/20 relation between training and testing sets: a pedagogical explanation," 2018. 54

[130] G. Y. Lee, L. Alzamil, B. Doskenov, and A. Termehchy, "A survey on data cleaning methods for improved machine learning model performance," *arXiv preprint arXiv:2109.07127*, 2021. 57

[131] E. Calisir and M. Brambilla, "The problem of data cleaning for knowledge extraction from social media," in *International Conference on Web Engineering*, pp. 115–125, Springer, 2018. 59

[132] A. Giachanou and F. Crestani, "Like it or not: A survey of twitter sentiment analysis methods," *ACM Computing Surveys (CSUR)*, vol. 49, no. 2, pp. 1–41, 2016. 60

[133] E. Colleoni, A. Rozza, and A. Arvidsson, "Echo chamber or public sphere? predicting political orientation and measuring political homophily in twitter using big data," *Journal of communication*, vol. 64, no. 2, pp. 317–332, 2014. 60

[134] G. Kesavaraj and S. Sukumaran, "A study on classification techniques in data mining," in *2013 fourth international conference on computing, communications and networking technologies (ICCCNT)*, pp. 1–7, IEEE, 2013. 60

[135] Z. Wang, Y. Bu, D. Bai, B. Wu, and J. Qin, "Bdap: A big data analysis platform based on spark," in *Journal of Physics: Conference Series*, p. 012023, IOP Publishing, 2018. 61

[136] J. Beel, B. Gipp, S. Langer, and C. Breitinger, "Paper recommender

systems: a literature survey," *International Journal on Digital Libraries*, vol. 17, no. 4, pp. 305–338, 2016. 62

[137] J. Hao and T. K. Ho, "Machine learning made easy: a review of scikit-learn package in python programming language," *Journal of Educational and Behavioral Statistics*, vol. 44, no. 3, pp. 348–361, 2019. 62, 63, 68, 94

[138] E. Bisong, *Building machine learning and deep learning models on Google cloud platform: A comprehensive guide for beginners.* Apress, 2019. 62

[139] E. Abdelzaher, "The systematic adaptation of violence contexts in the isis discourse: A contrastive corpus-based study," *Corpus Pragmatics*, vol. 3, no. 2, pp. 173–203, 2019. 79

[140] C. Hackett, B. Grim, M. Stonawski, V. Skirbekk, M. Potančoková, and G. Abel, "The global religious landscape," *Washington, DC: Pew Research Center*, 2012. 80

[141] N. Li and W. H. Murphy, "Religious affiliation, religiosity, and academic performance of university students: Campus life implications for us universities," *Religion & Education*, vol. 45, no. 1, pp. 1–22, 2018. 80

[142] Q. Li, S. Shah, R. Fang, A. Nourbakhsh, and X. Liu, "Tweet sentiment analysis by incorporating sentiment-specific word embedding and weighted text features," in *2016 IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, pp. 568–571, IEEE, 2016. 93

[143] O. Sharif, M. M. Hoque, A. Kayes, R. Nowrozy, and I. H. Sarker, "Detecting suspicious texts using machine learning techniques," *Applied Sciences*, vol. 10, no. 18, p. 6527, 2020. 93

[144] R. Z. Ul, S. Abbas, M. A. Khan, G. Mustafa, H. Fayyaz, M. Hanif, and M. A. Saeed, "Understanding the language of isis: An empirical approach to detect radical content on twitter using machine learning," 2021. 93

[145] J. Jamnani and M. Pandya, "Coordination of svc and tcsc for management of power flow by particle swarm optimization," *Energy Procedia*, vol. 156, pp. 321–326, 2019. 94

[146] M. Meissa, S. Benharzallah, L. Kahloul, and O. Kazar, "A personalized recommendation for web api discovery in social web of things.," *Int. Arab J. Inf. Technol.*, vol. 18, no. 3A, pp. 438–445, 2021. 96

[147] M. C. E. Meftah, "A new approach for social networks based on ontology of multilingual dynamic groups," *International Journal of Organizational and Collective Intelligence (IJOCI)*, vol. 12, no. 1, pp. 1–21, 2022. 99

[148] R. J. Turner, F. Coenen, F. Roelofs, K. Hagoort, A. Härmä, P. D. Grünwald, F. P. Velders, and F. E. Scheepers, "Information extraction from free text for aiding transdiagnostic psychiatry: constructing nlp pipelines tailored to clinicians' needs," *BMC psychiatry*, vol. 22, no. 1, pp. 1–11, 2022. 100

[149] A. Yeung, A. Iaboni, E. Rochon, M. Lavoie, C. Santiago, M. Yancheva, J. Novikova, M. Xu, J. Robin, L. D. Kaufman, *et al.*, "Correlating natural language processing and automated speech analysis with clinician assessment to quantify speech-language changes in mild cognitive impairment and alzheimer's dementia," *Alzheimer's research & therapy*, vol. 13, no. 1, pp. 1–10, 2021. 100

[150] S. Sun, C. Luo, and J. Chen, "A review of natural language processing techniques for opinion mining systems," *Information fusion*, vol. 36, pp. 10–25, 2017.

[151] H. Z. Brooks and M. A. Porter, "A model for the influence of media on the ideology of content in online social networks," *Physical Review Research*, vol. 2, no. 2, p. 023041, 2020.

[152] D. W. Franks, J. Noble, P. Kaufmann, and S. Stagl, "Extremism propagation in social networks with hubs," *Adaptive Behavior*, vol. 16, no. 4, pp. 264–274, 2008.

[153] S. Uddin, A. Khan, M. E. Hossain, and M. A. Moni, "Comparing different supervised machine learning algorithms for disease prediction," *BMC medical informatics and decision making*, vol. 19, no. 1, pp. 1–16, 2019.

[154] V. Chang, K. Hall, Q. A. Xu, Z. Wang, *et al.*, "A social network analysis of two networks: Adolescent school network and bitcoin trader network," *Decision Analytics Journal*, p. 100065, 2022.

[155] F. L. Tibbitts, "Deliberative democratic decision making, universal values, and cultural pluralism: A proposed contribution to the prevention of violent extremism through education," *Prospects*, vol. 48, no. 1, pp. 79–94, 2020.

[156] B. Qublan Al-Osaimi and Badr, "Intellectual extremism its definition, causes, manifestations, effects and ways to eliminate it," *Journal of the College of Education. banha*, vol. 29, no. 115 July C 1, pp. 1–15, 2018.

[157] J. Torregrosa, J. Thorburn, R. Lara-Cabrera, D. Camacho, and H. M. Trujillo, "Linguistic analysis of pro-isis users on twitter," *Behavioral Sciences of Terrorism and Political Aggression*, vol. 12, no. 3, pp. 171–185, 2020.

[158] A. Berhoum, M. C. E. Meftah, and S. Gourzi, "An intelligent approach for educational levels classification in social networks," *International Journal on Islamic Applications In Computer Science And Technology*, vol. 9, no. 2, pp. 1–19, 2021.