Alkadhum Journal of Science, Vol. 1, No. 2 (2023)

Open Access ISSN: 3005-2386 (Online), ISSN: 3005-2378 (Print)

IRAQI Academic Scientific Journals

Alkadhum Journal of Science (AKJS)

Journal Homepage: https://alkadhum-col.edu.iq/JKCEAS

AKJS Alkadhum Journal of Science

Opinion Mining in Arabic Extremism Texts: A Systematic Literature Review

[1]Ali Abbas Hadi Al-Shukrawi, [2]Layla Safwat Jamil, [3]Israa Akram Alzuabidi, [4]Ahmed Salman Al-Gamal , [5]Shahrul Azman Mohd Noah, [6]Mohammed Kamrul Hasan, , [7]Sumaia Mohammed Al-Ghuribi , [8]Rabiu Aliyu, [9]Zainab Kadhim Jabal and [10]Amjed Abbas Ahmed*

[1,4,9,10] Imam Al-Kadhum College (IKC) , Baghdad, Iraq
[2]College of Agricultural Engineering Sciences, University of Baghdad, Baghdad, Iraq
[3]Continuing Education Unit, College of Arts, University of Baghdad, Baghdad, Iraq
[5,6,7,8]Center for Cyber Security, Faculty of Information Science and Technology, University Kebangsaan Malaysia (UKM),Bangi 43600, Malaysia

Article information

Article history:
Received: October, 10, 2023
Accepted: November, 11, 2023
Available online: December, 14, 2023

Keywords:
Analysis,
Arabic text,
Extremism,
Opinion mining techniques,
Sentiment

Corresponding Author:
Amjed Abbas Ahmed
amjedabbas@alkadhum-col.edu.iq

DOI:
https://doi.org/10.61710/akjs.v1i2.60

Abstract

In this paper, a systematic literature review was provided that investigated the present evidence regarding extremist words in Arabic opinion mining methods. This study aimed to perform a Systematic Literature Review (SLR) in order to detect, evaluate, and synthesize the existing evidence regarding opinion mining techniques for extremist Arabic text. From the SLR, it is evident that opinion-mining techniques have several opportunities for detecting extremism in the Arabic text. Over the past few years, multimedia sentiment analysis has gained traction as visual content is becoming more incorporated into social media networking. Opinion mining is the process of identifying, extracting, and categorizing views about anything. It is a sort of Natural Language Processing (NLP) used to track public sentiment about a certain law, policy, or marketing, for example. It entails the creation of a method for collecting and analyzing comments and opinions concerning legislation, regulations, policies, and so on that are posted on social media. The process of information extraction is critical since it is both a beneficial tool and a difficult undertaking. In this article, we have examined the recent and advanced methodologies to extract sentiment from a web-wide item, opinion-mining methods must be automated. Also, we have analyzed the novel Artificial Intelligence and lexical-based algorithms for sentiment analysis. These methodologies find better applications in the customer feedback analysis of any organization.

1. Introduction

In the current era, people largely utilize the Internet to nowadays, the internet to gain access to various kinds of information and data [1]. People constantly send, receive, and retrieve data. Text-based content dominates the internet. In the 21st century, the information revolution is a major technological change. Globalization has resulted

from this transformation. Internet sites, particularly social media sites, are growing faster than expected. Also, this technical innovation has enabled everyone to express their thoughts on these internet platforms using text, visuals, and drawings. The internet is now used in every aspect of life, including politics, business, and society. This has resulted in a significant volume of internet data, which many scholars are interested in. Researchers' increasing interest may be attributed to the internet's many chances to gather knowledge and apply it to personal interests [2]. Opinion mining is often linked to business intelligence. Many studies have mined customer feedback to discover market trends. Importantly, government intelligence uses opinion mining. This partnership highlights popular opinion on important matters including elections, policy, manifestos, etc. Customer reviews strongly influenced online purchases, according to Pang et al. (Yu and Kak, 2012). Online reviews help readers choose products [3]. Most internet review research focuses on movies and corporate items, the authors discovered. Extremist political opinion mining is rare and undesirable. In this study, we have focused on the extremist content in the Arabic language mainly because of the Arab Spring that erupted during the previous decade [4]. Online publications, news, and views covered more terrorist activities in Arab nations. Consequently, this emphasizes the necessity to analyses and examine such information and articles to determine the decision-making flaws that led to extremism. Opinion mining should identify Arabic extreme text using them. It prompted experts to examine these incidents and how extremism affects individuals [5].

It has been seen that most of the research studies on opinion mining focus on consumer reviews [6]. Yet, extremist material research is scarce, and media experts, public stakeholders, and the general public are concerned about it. Print and internet radicalism are merging. News websites routinely update extreme stories and viewpoints. This work discusses data mining extremism detection. Opinion mining was used here. Analysts may easily examine and analyze extremist articles online if they write them. As anybody may publish extremist information online, there are many sources. As this is unstructured data, finding and classifying extremist articles is important. This categorization helps decision-makers make informed choices to prevent escalation. Extremism researchers are mainly concerned with public opinions regarding the government and leaders [7]. Academic extremism analysts also recognize these beliefs, but they focus on countering extreme situations. Public intelligence services investigate public opinions on government policy to predict popular responses [8]. Moreover, the vote bank of a political party is significantly affected by extremism-related policies and manifestos as people really inclined toward this special issue [9]. In opinion mining, extremism detection entails two value classifications: extremism and non-extremism.

In this study, the following research question has been addressed: What opinion-mining techniques are employed for the classification of extremism in Arabic content? As conformance with this research question, the main objective of this study was to undertake a systematic literature review that evaluates the opinion mining techniques that are employed to detect extremism in Arabic content. In general, this article aims to provide a complete account of opinion- mining methods used to detect extremism. This paper is divided into five main sections. The first section entails the background of the study. In the second section, methodology used to address the research question is highlighted. In the third section, the findings of the Systematic Literature Review (SLR) are illustrated the then they are discussed in related with previous literature in the fourth section. In the fifth and final section, the study presents its conclusions.

2. Methods

2.1 Search methods

Findings of the (Systematic Literature Review) SLR are illustrated and then they are discussed concerning previous literature in the fourth section. In the fifth and final section, the study presents its conclusions. To undertake the search, five databases were investigated in total to find research papers related to opinion mining approaches for extremist Arabic text. These included Google Scholar, Scopus, Springer ProQuest, and EBSCOhost are among the databases considered. In order to extract the most relevant articles from these databases, a well-structured search technique was devised. Boolean logic was utilized to restrict and widen the search using the phrases "AND" and "OR" The following two search terms were utilized in particular:

- (Opinion mining techniques OR methods AND challenges AND extremism OR extremism Arabic text).
- (Extremism OR extremism Arabic text AND risks OR threats AND opinion mining techniques OR methods).

2.2 Eligibility Criteria

The following rules were followed to be included in the review:

- In order to evaluate meta-synthesis-related literature, an academic publication based on a qualitative methodological approach was required. As a result, books and grey literature were not included in this literature.
- The opinion-mining techniques employed in the research study should iterate on general extremism or extremism in Arabic content.
- It must be ensured that the guaranteed timeliness and relevancy should be published between 2016-2021.
- The research has to be written in English and published in the corresponding language.

2.3 Extraction Method and Synthesis

The researcher identified twelve eligible papers which justified the criteria for selection. From these papers, necessary data was retrieved. In order to pinpoint the themes, present in the studies, a qualitative thematic analysis was undertaken. From these studies, the following sections were examined for analysis: purpose of research and objectives, research approach, findings, practical and theoretical significance, and limitations. The data that was extracted from these studies were subjected to extensive collation, summarization, and categorization. Then, the data related to the interest of the research were taken from the data. After this, the collected data were structured and organized to change it from raw form into refined form. The findings of the Systematic Literature Review (SLR) were summarized in the tabulated form.

2.4 Quality Criteria

The quality evaluation is a pivotal step in SLR as it helps in ensure that the review entails competent and peer-reviewed studies and shows concordance with the research objectives. To enhance the reliability and validity of the research findings, the empirical quality of the studies should be subjected to extensive quality review [10]. To achieve such quality, the following questions were taken into consideration during the quality assessment process:

1. Does the name(s) of the author(s) is provided in clear wording?
2. Does the aims and objectives of the study are explicitly defined?
3. Does the study entail an organized and viable research methodology to address the research question(s).
4. Does the study explicitly describe its results and discussions?
5. Is the research paper published in a reputable journal?

3.Findings

3.1 Study Flow

More than 8,054 articles were returned after inserting the keywords into the relevant academic databases. Journal articles that were replicated across several databases were excluded from the final review. The evaluation allowed the authors to filter 7998 research papers based on various features such as research titles. Moreover, the papers were also subjected to screening by employing the abstract and full text, which resulted in the elimination of 41 studies. Based on this screening process, a total of 12 research papers were selected that justified the predefined inclusion and quality criteria.

3.2 Attributes of Including Studies

Table 1 highlights the features of the studies that were included. Authors and publication years are prioritized, as are aim(s) and/or objectives and research techniques.

Table (1): Attributes of included studies.

Authors and publication years	Aims and/or objectives	Research methods
[11]	To give an overview of sentiment analysis (SA) and opinion mining (OM) methodologies, as well as the many techniques employed in this discipline. It also discusses sentiment analysis applications, challenges, and prior research.	A literature review.
[12]	To perform an evaluation of existing methodologies and techniques for detecting, analyzing, and predicting human behaviors by	A structured literature review.

	mining a range of textual data sources, with a focus on enabling the categorization of psychological behaviors related to emotion, cognition, and social empathy.	
[13]	To present a method for categorizing Arabic text based on word weighting and the rough set theory's reduction idea to decrease the number of terms necessary to produce the classification rules that comprise the classifier.	Upgraded Quick reduct creates a multiple minimal reduct extraction algorithm. The rough set classifier's classification rules use several reducts. The approach is tested on 2700 Arabic texts in nine categories. The experiment tested the approach with multiple and single minimal reducts.
Asif et al. [14]	To investigate the emotive analysis of multilingual textual data from social media to determine the severity of extremist attitudes. The combined textual viewpoints are also categorized by extremism: high, low, moderate, and neutral.	Starting with an intensity-weighted multilingual lexicon. Domain experts validated this lexicon at 88%. Classification follows using Multinomial Nave Bayes and Linear Support Vector Classifier.
[14]	To present a system that would enable to pinpointing of significant change-points in time series relevant to extremism that could help in determining the occurrence of activities that require increasing monitoring	The suggested methodology entails: (i) the collected textual data were divided into categories based on terrorism and hate speech. These may serve as the representation of the critical activities; and (ii) to conduct a thorough analysis of the change-points present in the textual data. The technique was evaluated using a publically accessible jihadist forum dataset. The last process identifies the topic using predicted change points to improve accuracy.
[15]	To present a system that would enable to pinpointing of significant change points in time series relevant to extremism that could help in determining the occurrence of activities that require increasing monitoring.	The suggested methodology entails: (i) the collected textual data were divided into categories based on terrorism and hate speech. These may serve as the representation of the critical activities; and (ii) to conduct a thorough analysis of the change-points present in the textual data. A publicly accessible dataset from jihadist forums was used to assess the method. In the last phase, the topic is selected using predicted change points to test accuracy.
[16]	To devise a machine learning-based approach to help identify tweets imbued with extremist comments. The study also analyzes the advantages and disadvantages of using principal component analysis as an exploratory data analysis in terms of textual data which is usually with a large number of factors.	EDA and principal component analysis reduced Twitter data to low-dimensional space. The work used PCA-based reduced characteristics from extremism-related tweets to train multiple ML classification algorithms (KNN, SVM, etc.).
[17]	To offer a paradigm for analyzing terrorism-related information, with an emphasis on categorizing tweets as extremist or non-extremist.	Extremist and non-extremist tweets were classified using deep learning. The data were

		additionally tokenized, stop word removed, etc.
[18]	To identify methods for automatically detecting radical material on social media.	The research used a two-part technique. The first section used Dabiq extremist publications to build a linguistic model. TF-IDF scores of unigrams, bigrams, and trigrams and word2vec model word embeddings are given into this phase to construct a language model. Phase 2: Tweet classification creates radical activity characteristics using Phase 1 models.
[19]	To examine the comprehension of users regarding extremism in Arabic and how they use this word in daily life to build and transfer their opinion regarding the term.	Several tweets were gathered from Twitter using the Arabic phrase "تطرف أو متطرف" meaning extremist/extremism. Topic modeling was used to identify latent extremism-related topics. Van Dijk's Socio-cognitive technique was used to analyze critical dialogue. Users' discursive and linguistic strategies to justify extremism were examined.
[20]	To assess the current status of text mining research by reviewing advances in published literature over the last several years, and to give significant insights for practitioners and scholars on the main trends, methodologies, and applications of text mining research.	A structured literature reviews.
[21]	To evaluate the issue of opinion mining of reviews and comments written in colloquial Arabic. Moreover, it also aimed to assess the ability of ML algorithms.	Dialect stemmers are lightweight. Two feature sets are tested for sentiment analysis. Lastly, Nave-Bayes, SVM, and Maximum Entropy ML opinion mining performance are examined. F1-measures ML algorithm performance.
[22]	To provide an overview of the issue of opinion mining and to present the algorithms for opinion mining for networking sites.	A comprehensive questionnaire of the current literature yielded a complete assessment of visual sentiment analysis.

3.3 Synthesis of Data

The systematic literature review produced indications on opinion mining methodologies for extremists in general and in Arabic text in particular. Table 2 summarizes the Systematic Literature Review (SLR) findings by the primary themes visible in the chosen publications. To characterize diverse opinion mining methodologies, the research used distinct terminology. As a result, the terminologies were standardized in order to establish standardization and ease of analysis.

Table (2): Themes related to opinion mining approaches.

No.	Themes related to opinion mining approaches	Brief description	Source
1	Sentiment Analysis	Sentiment analysis may be used for the extraction of people's thoughts and attitudes inherent in user-constructed information, allowing for a broad range of implementations such as political vote predictions, and public opinion monitoring. Nowadays, sentiment analysis for social media multimedia material is crucial. As it is crucial to planning, thinking, and decision-making.	[11], [12], [14], [17], [22]

2	Extremism	Extremism has many contradicting meanings, making it hard to describe. Extremism is unusual views, attitudes, emotions, actions, and methods. Extremism happens when a person doesn't allow for other viewpoints, considers their own beliefs exclusive, and doesn't allow for variety. Extremism also happens when someone uses violence to impose their views on others.	[15], [16], [18],
3	Lexicon based approaches	Form en uses emotion lexicons like "attacked" for harsh moods and "happy" for mild moods. Lexical polarity may vary by domain. Using a lexical dictionary, this approach determines sentiment orientation. This approach helps analyze document and phrase text.	[14], [16], [17], [21], [22]
4	Machine learning-based approaches	Machine learning is referred to as learning through experience and data in which the information systems are trained again and again until they become automatic. ML is considered one of the most vital parts of AI. In ML, a sample or training dataset is utilized to build a model by using compatible algorithms. It helps to predict from the datasets with prior programing. ML algorithms are useful in medicine, email filtering, voice recognition, and computer vision. ML algorithms outperform classical algorithms.	[11], [13], [16], [17], [18], [20], [21]
5	Deep learning-based approaches	ML-based deep learning uses ANNs and representation learning. Semi-supervised or unsupervised learning is possible. Deep-learning techniques like deep neural networks and convolutional neural networks are used in computer vision, machine translation, bioinformatics, medical image analysis, and others.	[12], [15], [17], [22]
6	Text Categorization	Text categorization assigns a category to an unlabeled document based on its content. Automatic text classification may be hindered by the data's high dimensionality.	[13]

4. Discussion

4.1 RE-Examining the Research Question

The purpose of this study based on Systematic Literature Review (SLR) was to investigate opinion mining strategies for extremist Arabic text. Opinion mining encompasses a series of methodologies that are part of a larger discipline referred to as text mining. This mining technique is widely regarded as the sub-discipline of the data mining technique. Research on opinion mining began during the 2000s although the terminology of "opinion mining" is associated with [23]. Much research has been undertaken over the last 18 years to explore and assess the viewpoints conveyed in articles, news, product, and service evaluations.

4.2 Previous Studies

Researchers have attempted to define opinion mining. It's a kind of text mining that extracts valuable data from user text. Its usage on websites is rather recent [24]. Newspapers, opinion blogs, television sites, and social media have increased its vibrancy throughout the years. Opinion mining has revolutionized Natural Language Processing (NLP) [25]. This mining approach detects online user views on the subject of interest. This method lets the researcher assess text-based user emotions and attitudes. They may then analyze the data to derive user views to better understand the problem [26].

Opinion mining extracts people's views on a document's topic. The object may be a mood, product, service, concept, idea, news policy, extremism, people, or other noteworthy event. After recognizing the item, this method extracts key characteristics from textual data. Positive, negative, or neutral remarks are assessed [23]. This mining method relies on subjectivity. Opinion mining evaluates things of interest subjectively.

The term opinion mining has been employed by the researchers under various synonymous terms such as sentiment analysis [27], sentiment classification [28], affect analysis [29], and opinion analysis [30]. Opinion extraction, emotion mining, review mining, and many more are similar phrases [26]. Opinion mining now encompasses these concepts. Industry uses sentiment analysis more often. Yet, academics utilize sentiment analysis and opinion mining. Opinion mining is increasingly common in student theses and dissertations.

The term "opinion mining" is largely associated with [23]. He suggested strategies for categorizing views as good or unfavorable. This opinion mining method categorized textual data by user emotions, views, and sentiments.

This method includes automated user opinion extraction from textual data [31]. Opinion mining differs from standard data mining in many ways. This data mining method involves imagination to categorize data, which is difficult [32]. For instance, classification by word count of positive and negative terms are difficult since comparable words have various meanings, making it difficult to create such a list [33].
Text mining or text data mining [34] can be simply referred to as a data mining technique used to extract knowledge from textual data [35]. It also refers to NLP's method of extracting important facts from raw or unstructured input [36, 37]. Opinion mining is challenging because raw data is hard to manage and structure[38]. Algorithms can extract essential information from raw natural language data to solve this challenge [39]. Machines cannot grasp raw data as well as people do. Hence, textual data may provide vital information [40]. Text mining organizes raw data. Patterns are developed to organize raw data in this method. At the last stage, data is analyzed [41].
The Text mining approach [35] can be described as the process through which structured text such as RDBMS is utilized and valuable knowledge is retrieved [40, 42]. Moreover, semi-structured data such as XML [43], and unstructured textual data such as word documents, web pages, and comments are also utilized to extract critical information [44]. Text mining covers a lot of important data and uses numerous methods to evaluate textual data in Arabic extremist messages. These writings involve information retrieval (IR) communities [45, 46], NLP [47, 48], information extraction [49, 50], text summary [51], machine learning [52], biomedical sciences [53], text categorization [54, 55], and sentiment analysis [4].
Arabic is complex. Stemming is necessary for text mining this language's morphological structure. The Arabic language entails these stemming processes including [56, 57] root stemming [58] and light stemming [59]. Arabic uses stemming procedures differently from other languages. Room stemming determines Arabic word roots morphologically [55]. The two most effective stemmers for Arabic root stemming are the Information Science Research Institute [60] and Khoja [59].
In order to remove redundant prefixes, light stemming is generally employed [59, 61, 62]. Arabic light stemming is unstandardized. Word meaning changes distinguish the two stemming procedures [63]. For instance, the word "الثورة" which means revolution, stems from "ثور" meaning ox. Most root stemmers, including ISRI and Khoja, remove prefixes, infixes, and suffixes, changing the meaning. However, in case of light stemming, it would be used as "ثورة", which results in a change in the meaning of the word as compared to other roots.

5.Conclusion

The threat of extremism to academic institutions and national security is one of the most pressing problems studied by academics. Opinions are extremely important in extremism. Extremism is a considerably more difficult issue to study owing to its complexities and emotional manifestations. Nowadays, the internet plays an important role in allowing the general public to read radical viewpoints and participate in political conversations about various political events. As a result, it is advantageous to study opinions in Arabic extremism postings in order to comprehend extremism detection. Extremism opinion mining may be used in a variety of domains to achieve a variety of purposes. The advantage of extremist opinion mining above standard surveys is that it provides information on "why individuals support or oppose specific government policies." The majority of news portals allow visitors to share their thoughts on political matters. Furthermore, they supply government securities with what is released and utilized for everything government-related. Education is extremely significant for a college of political sciences since it provides more knowledge about the issue, allows a student to deal with that subject, and puts a wealth of material at the disposal of both students and professors. Most institutions now provide academic modules that require students to provide online feedback on their courses, team members, and instructors. Therefore, future studies should increase their focus on empirical evaluation of opinion mining techniques to reach a more robust understanding of these techniques for extremism in Arabic text. In this article, we have analyzed the recent and advanced methodologies to extract sentiment from a web-wide item, opinion-mining methods must be automated. Finally, we have presented a detailed review of the novel Artificial Intelligence and lexical-based algorithms for sentiment analysis. These methodologies are well-suitable for applications in the customer feedback analysis of any organization.

Funding

NONE

Acknowledgments

The Authors would like to thank Imam AL-Kadhum College (IKC) (https://alkadhum-col.edu.iq/) in Baghdad - Iraq for its support in the present work.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] M. Graham and W. H. Dutton, Society, and the internet: How networks of information and communication are changing our lives. Oxford University Press, 2019.

[2] U. Can and B. Alatas, "A new direction in social network analysis: Online social network analysis problems and applications," Physica A: Statistical Mechanics its Applications, vol. 535, p. 122372, 2019.

[3] S. M. Al-Ghuribi and S. A. M. Noah, "Multi-criteria review-based recommender system–the state of the art," IEEE Access, vol. 7, pp. 169446-169468, 2019.

[4] B. Pang and L. Lee, "Opinion mining and sentiment analysis," Foundations and Trends® in Information Retrieval, vol. 2, no. 1–2, pp. 1-135, 2008.

[5] A. A. Ahmed et al., "Extremism Arabic Text Detection using Rough Set Theory: Designing a Novel Approach," IEEE ACCESS, 2023.

[6] G. Wolfsfeld, E. Segev, and T. Sheafer, "Social media and the Arab Spring: Politics comes first," The International Journal of Press/Politics, vol. 18, no. 2, pp. 115-137, 2013.

[7] P. Chaovalit and L. Zhou, "Movie review mining: A comparison between supervised and unsupervised classification approaches," in Proceedings of the 38th annual Hawaii International Conference on System Sciences, 2005, pp. 112c-112c: IEEE.

[8] D. H. Abd, A. T. Sadiq, and A. R. Abbas, "Classifying Political Arabic Articles Using Support Vector Machine with Different Feature Extraction," in International Conference on Applied Computing to Support Industry: Innovation and Technology, 2019, pp. 79-94: Springer.

[9] G. Stylios et al., "Public opinion mining for governmental decisions," Electronic Journal of e-Government, vol. 8, no. 2, p. 202, 2010.

[10] M. Petticrew and H. Roberts, Systematic reviews in the social sciences: A practical guide. John Wiley & Sons, 2008.

[11] B. Saberi and S. Saad, "Sentiment analysis or opinion mining: a review," Int. J. Adv. Sci. Eng. Inf. Technol, vol. 7, no. 5, pp. 1660-1666, 2017.

[12] E. Gutierrez, W. Karwowski, K. Fiok, M. R. Davahli, T. Liciaga, and T. J. S. Ahram, "Analysis of human behavior by mining textual data: current research topics and analytical techniques," vol. 13, no. 7, p. 1276, 2021.

[13] Q. A. Al-Radaideh and M. A. Al-Abrat, "An Arabic text categorization approach using term weighting and multiple reducts," Soft Computing, vol. 23, no. 14, pp. 5849-5863, 2019.

[14] M. Asif, A. Ishtiaq, H. Ahmad, H. Aljuaid, and J. Shah, "Sentiment analysis of extremism in social media from textual information," Telematics Informatics, vol. 48, p. 101345, 2020.

[15] O. Theodosiadou et al., "Change point detection in terrorism-related online content using deep learning derived indicators," vol. 12, no. 7, p. 274, 2021.

[16] W. Sharif et al., "An empirical approach for extreme behavior identification through tweets using machine learning," vol. 9, no. 18, p. 3723, 2019.

[17] S. Ahmad, M. Z. Asghar, F. M. Alotaibi, and I. Awan, "Detection and classification of social media-based extremist affiliations using sentiment analysis techniques," Human-centric Computing Information Sciences, vol. 9, no. 1, pp. 1-23, 2019.

[18] M. Nouh, J. R. Nurse, and M. Goldsmith, "Understanding the radical mind: Identifying signals to detect extremist content on Twitter," in 2019 IEEE International Conference on Intelligence and Security Informatics (ISI), 2019, pp. 98-103: IEEE.

[19] S. A. Hamdi, "Mining ideological discourse on Twitter: The case of extremism in Arabic," Discourse Communications of the ACM, vol. 16, no. 1, pp. 76-92, 2022.

[20] H. Hassani, C. Beneki, S. Unger, M. T. Mazinani, and M. R. Yeganegi, "Text mining in big data analytics," Big Data Cognitive Computing, vol. 4, no. 1, p. 1, 2020.

[21] A. Y. Al-Obaidi and V. W. Samawi, "Opinion mining: analysis of comments written in Arabic colloquial," in Proceedings of the World Congress on Engineering and Computer Science, 2016, vol. 1.

[22] Z. Li, Y. Fan, B. Jiang, T. Lei, and W. Liu, "A survey on sentiment analysis and opinion mining for social multimedia," Multimedia Tools Applications, vol. 78, no. 6, pp. 6939-6967, 2019.

[23] K. Dave, S. Lawrence, and D. M. Pennock, "Mining the peanut gallery: Opinion extraction and semantic classification of product reviews," in Proceedings of the 12th International Conference on World Wide Web, 2003, pp. 519-528.
[24] S. Bhatia, P. Chaudhary, and N. Dey, "Opinion Summarization," in Opinion Mining in Information Retrieval: Springer, 2020, pp. 81-95.
[25] P. Balaji, D. Haritha, and O. Nagaraju, "An overview on opinion mining techniques and sentiment analysis," International Journal of Pure and Applied Mathematics, vol. 118, no. 19, pp. 61-69, 2018.
[26] M. Z. Asghar, "Opinion Extraction From Online Blogs And Public Reviews," GOMAL UNIVERSITY DI KHAN, 2014.
[27] D. K. Evans, L.-W. Ku, Y. Seki, H.-H. Chen, and N. Kando, "Opinion analysis across languages: An overview of and observations from the NTCIR6 opinion analysis pilot task," in International Workshop on Fuzzy Logic and Applications, 2007, pp. 456-463: Springer.
[28] Z. Fei, J. Liu, and G. Wu, "Sentiment classification using phrase patterns," in The Fourth International Conference on Computer and Information Technology, 2004. CIT'04., 2004, pp. 1147-1152: IEEE.
[29] P. Subasic and A. Huettner, "Affect analysis of text using fuzzy semantic typing," IEEE Transactions on Fuzzy Systems, vol. 9, no. 4, pp. 483-496, 2001.
[30] E. Riloff, S. Patwardhan, and J. Wiebe, "Feature subsumption for opinion analysis," in Proceedings of the 2006 conference on empirical methods in natural language processing, 2006, pp. 440-448.
[31] F. Hemmatian and M. K. Sohrabi, "A survey on classification techniques for opinion mining and sentiment analysis," Artificial Intelligence Review, pp. 1-51, 2017.
[32] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up?: sentiment classification using machine learning techniques," in Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10, 2002, pp. 79-86: Association for Computational Linguistics.
[33] D. C. Dias, "Text mining methods for mapping opinions from georeferenced documents," 2012.
[34] M. A. Hearst, "Text data mining: Issues, techniques, and the relationship to information access," in Presentation notes for UW/MS workshop on data mining, 1997, pp. 112-117.
[35] R. Feldman and I. Dagan, "Knowledge Discovery in Textual Databases (KDT)," in KDD, 1995, vol. 95, pp. 112-117.
[36] T. A. Al-asadi, A. J. Obaid, R. Hidayat, and A. A. Ramli, "A survey on web mining techniques and applications," International Journal on Advanced Science Engineering and Information Technology, vol. 7, pp. 1178-1184, 2017.
[37] M. P. Singh, The practical handbook of Internet computing. CRC Press, 2004.
[38] S. M. Al-Ghuribi and S. Alshomrani, "A simple study of webpage text classification algorithms for Arabic and English Languages," in 2013 International Conference on IT Convergence and Security (ICITCS), 2013, pp. 1-5: IEEE.
[39] A. Akilan, "Text mining: Challenges and future directions," in 2015 2nd International Conference on Electronics and Communication Systems (ICECS), 2015, pp. 1679-1684: IEEE.
[40] M. Allahyari et al., "A brief survey of text mining: Classification, clustering and extraction techniques," arXiv preprint arXiv:1707.02919, 2017.
[41] M. K. Saad and W. M. Ashour, "Arabic text classification using decision trees," Arabic text classification using decision trees, vol. 2, 2010.
[42] M.-S. Chen, J. Han, and P. S. Yu, "Data mining: an overview from a database perspective," IEEE Transactions on Knowledge and Data Engineering, vol. 8, no. 6, pp. 866-883, 1996.
[43] S. A. Pouriyeh, M. Doroodchi, and M. Rezaeinejad, "Secure Mobile Approaches Using Web Services," in SWWS, 2010, pp. 75-78.
[44] S. M. Al-Ghuribi, S. A. M. Noah, and S. Tiun, "Unsupervised Semantic Approach of Aspect-Based Sentiment Analysis for Large-Scale User Reviews," IEEE Access, vol. 8, pp. 218592-218613, 2020.
[45] C. Faloutsos and D. W. Oard, "A survey of information retrieval and filtering methods," 1998.
[46] C. D. Manning, P. Raghavan, and H. Schütze, Introduction to information retrieval. Cambridge University Press, 2008.
[47] E. D. Liddy, "Natural language processing," 2001.
[48] C. D. Manning, C. D. Manning, and H. Schütze, Foundations of statistical natural language processing. MIT press, 1999.

[49] J. Cowie and W. Lehnert, "Information extraction," Communications of the ACM, vol. 39, no. 1, pp. 80-91, 1996.
[50] T. Mitsumori, S. Fation, M. Murata, K. Doi, and H. Doi, "Gene/protein name recognition based on support vector machine using dictionary as features," BMC bioinformatics, vol. 6, no. S1, p. S8, 2005.
[51] D. R. Radev, E. Hovy, and K. McKeown, "Introduction to the special issue on summarization," Computational linguistics, vol. 28, no. 4, pp. 399-408, 2002.
[52] C. C. Aggarwal and C. Zhai, Mining text data. Springer Science & Business Media, 2012.
[53] J. B. Gutierrez, M. R. Galinski, S. Cantrell, and E. O. Voit, "From within-host dynamics to the epidemiology of infectious disease: Scientific overview and challenges," ed: Elsevier, 2015.
[54] S. Alsaleem, "Automated Arabic Text Categorization Using SVM and NB," Int. Arab J. e-Technol., vol. 2, no. 2, pp. 124-128, 2011.
[55] A. Noaman and S. Al-Ghuribi, "A NEW APPROACH FOR ARABIC TEXT CLASSIFICATION USING LIGHT STEMMER AND PROBABILITIES," International Journal of Academic Research, vol. 4, no. 3, 2012.
[56] M. A. Otair, "Comparative analysis of Arabic stemming algorithms," International Journal of Managing Information Technology, vol. 5, no. 2, pp. 1-13, 2013.
[57] R. Abooraig, S. Al-Zu'bi, T. Kanan, B. Hawashin, M. Al Ayoub, and I. Hmeidi, "Automatic categorization of Arabic articles based on their political orientation," Digital Investigation, vol. 25, pp. 24-41, 2018.
[58] S. Khoja and R. Garside, "Stemming Arabic text," Lancaster, UK, Computing Department, Lancaster University, 1999.
[59] L. S. Larkey, L. Ballesteros, and M. E. Connell, "Improving stemming for Arabic information retrieval: light stemming and co-occurrence analysis," in Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval, 2002, pp. 275-282.
[60] K. Taghva, R. Elkhoury, and J. Coombs, "Arabic stemming without a root dictionary," in International Conference on Information Technology: Coding and Computing (ITCC'05)-Volume II, 2005, vol. 1, pp. 152-157: IEEE.
[61] M. Aljlayl and O. Frieder, "On Arabic search: improving the retrieval effectiveness via a light stemming approach," in Proceedings of the Eleventh International Conference on Information and Knowledge Management, 2002, pp. 340-347.
[62] L. S. Larkey and M. E. Connell, "Arabic Information Retrieval at UMass," in Text REtrieval Conference, Gaithersburg, Maryland, 2001, pp. 562-570.
[63] M. A. Alhanjouri, "Pre-processing techniques for Arabic documents clustering," International Journal of Engineering Management Research, vol. 7, no. 2, 2017.