Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

Emotion Detection in Arabic Text in Social Media: A Brief Survey

Zakaria H.Ali [1]*, Hiba J.Aleqabie[2]

[1] Computer Science Department, Collage of Computer Science and Information Technology University of Kerbala, Karbala, Iraq. *Email: Zakariya.h@s.uokerbala.edu.iq*

[2] Artificial Intelligence Engineering, Collage of Information Technology Engineering, Al-Zahraa University for Women. Karbala, Iraq.

[2] Computer Science Department, Collage of Computer Science and Information Technology University of Kerbala, Karbala, Iraq. *Email: Hiba.jabbar@Alzahraa.edu.iq*

* *Zakariya.h@s.uokerbala.edu.iq*

https://doi.org/10.46649/fjiece.v3.2.27a.6.6.2024

Abstract. *Emotion detection in Arabic text in social media has gained significant attention due to the increasing use of social media platforms in the Arab world. It is a special case of affective computing that relies on natural language processing techniques to recognize emotions from written text. This article provides a comprehensive review of different approaches used in developing models for emotion detection in Arab textual data in social media. Social media have a substantial impact on societies and have contributed significantly to many fields (marketing, recommendation systems, etc.). Arabic language texts on social networking sites in all their forms (posts - comments - messages - any other form) express the personal feelings of the person concerned; if they are understood correctly, the understanding of the written text will be correct. Scientific research has recently increased in this field to determine people's feelings through their writing of texts. In addition, we have highlighted the utilization of deep learning architectures in the improvement of the performances of models for different types of deep learning models.*

Keywords: Emotion detection; Arabic texts; machine learning; social media; natural language processing.

1. INTRODUCTION

The excellent success of emotion detectors, including machine learning (ML) approaches, is based on the large data amounts of annotated publicly accessible resources and studies, which help to identify, train, validate, and check whether the data has been transported into numerous natural language (NLP) apps [1][2]. Due to the scarcity of documents, including the number of lexicons and annotated datasets, such methods cannot be easily provided for Arabic emotions analysis. Although the production of resources, such as labelling lexicons, is crucial for developing learning-based machine techniques to accomplish state-of-the-art outcomes, we are primarily focused on providing high-quality learning-based emotion detection results. Different emoticons are created to improve the detection results of emotions in Arabic. Arabic has a different script entirely, with its own grammar and lexicons.

Social media sites are some of the most important tools that are used today as an outlet for the latest news, opinions, complaints, and praises [3][4]. It is a known fact that social media platforms are sources that generate substantial amounts of users' opinions and data. Since humans possess emotions, likewise emotional content is generated in this textual data; posts contain high expressions of feelings. Emotion detection in such Arabic texts on social media is important and practical in a variety of situations, for instance, helping the production of smart automatic chatting systems, emotion prediction, marketing problems, and prediction of the trends of markets [5]. The primary use of emotion detection in social

media texts of English-originated countries has been to provide options and investigations in Arabic texts that have proven to have more desired results.

Different languages have been developed due to the fact that social media, as a means of communication, is used frequently around the world. Thus, studies on emotion detection in social media are usually conducted in languages spoken in different regions of the world. There is a study in Persian for the language spoken in Iran and some countries in the outer vicinity, which comes from the studies of emotion detection based on lexical method, machine learning method, hybrid method, and deep learning method [6]. Although the English language is used objectively because it is widely known, many people, as much as it is spoken in the world due to the differences in common conversation, such as dialect and slang that language affects from the region, is rarely as a result of the similarity seen in all speakers all over the world. Emotion detection in English and German has meanwhile been made for languages such as compact between studies.

It is difficult to detect the specifics of the emotions that each person is experiencing. On the other hand, emotions are a very important part of human communication. Hence, it is vital to communicate more efficiently and try to understand the emotion. One of the advantages of understanding emotions at a very high level is that the information on the person can be expected to be inferred from who writes the text[1] . From this point, social media, due to its widespread use, is among one of the most important areas of communication used for emotion detection in today's communication medium [7].
The customer's text is of huge interest to natural language processing (NLP) engineers by tackling the emotional implications of their text. Emotion recognition in speech is said to be "the most modern work of the future," a data-driven approach. Relaxation application is another NLP research field to be interested in the behavioral science and social psychology of the shopper[2]. Rather than humans, detecting whether the text contents make consumers laugh, weep, or be furious. Typically, since the text is submitted electronically, it goes through a number of decentralization structures and devices, and bills are first received and labeled with any relevant data. This period has to be conducted so as to raise the accuracy of the system for information discovery in which there is a potential emotional recognition related to their target.
Social media contributes to the flood of data generated on a daily basis that rapidly spreads local and global knowledge to end-users [8]. Many social media platforms such as Twitter, Facebook, and Instagram provide users with an opportunity to type messages, among other media. It is in this type of social media communication that better emotional and psychological situations of people are available, unlike any other way of communication[9] [4], [10]. On a single day, more than 3.5 million text messages are submitted by practitioners in the United Arab Emirates (UAE), and hundreds of different staff access social media and other methods of digital communication. Most of the very poor people, however, address some emotional implications and show them through typing by psychological stress between at least one in these text communications [11].

2.Related works

A range of studies have explored emotion detection in Arabic text on social media platforms using machine learning. The researcher in[12] applied Word2Vec for word embedding and an ensemble XGBoost model to achieve high accuracy in sentiment analysis of Arabic tweets. The top-performing model was achieved through the utilization of hyperparameter optimization and Grid Search technology. However, the sample size of 41,125 tweets, although not insignificant, may be inadequate for making sweeping generalizations about sentiment detection in Arabic text. In[13] found that deep learning techniques, particularly with Word Embedding, outperformed machine learning in sentiment analysis of Arabic tweets about Saudi telecommunication companies where Deep learning technique with Word Embedding showed promising accuracy in sentiment analysis (F1=0.81). Rather than, Logistic Regression (LR), Support Vector Machine (SVM), and Random Forest (RF) using unigram language model outperformed the bigram language

representation in sentiment analysis, The article discusses a deficiency in current research by incorporating NLP methods such as POS tagging and Named Entity Recognition in the suggested frameworks. The assessment criteria utilized are confined to accuracy and F1 score, and a more extensive assessment encompassing precision, recall, and F1 score for each category could have offered greater insights. Moreover, The deep Convolutional Neural Network (CNN) was utilized by [14] to categorize emotions in Arabic tweets. The results of this approach were found to be superior to those achieved by traditional machine learning algorithms. This deep learning model follows a four-step process involving word, sentence, and document vectorization, ultimately leading to classification. This comprehensive approach appears to be quite effective in its end-to-end architecture. The assessment of the SemiEval dataset demonstrated the deep learning model's efficacy in analyzing emotions in Arabic text. However, the paper does not include a comprehensive error analysis to identify which types of emotions the deep learning model has difficulty classifying accurately compared to traditional machine learning approaches.
The paper does not provide any insights or analysis on the interpretability of the deep learning model and how the learned representations capture the nuances of emotional expressions in Arabic text. In [15]The ensemble classifiers with 10-fold cross-validation outperformed other machine learning classifiers in accuracy, specificity, precision, F1 score, and sensitivity. The research addresses the challenges of sentiment analysis for the Arabic language due to its complex morphology, orthography, and dialects. The paper lacks details on the specific ensemble techniques used, the dataset characteristics, and the feature engineering/preprocessing steps.

3. Challenges and Opportunities

Thus, one of the eagerly researched areas is the creative from a diversity of genres, including social and religious divergent datasets and blogs such as the Sadness, Fear, Joy, Anger, Disgust and Surprise (S.F.J.A.D.S.) theory of basic divergent workload geographical splits of Arabic regions and time specialty sciences, Arabic content is the main data generator for reflecting the full nature in the comment section of such films-text analysis [16][2].
Emotion detection models in the Arabic language face several challenges, including limited labeled datasets, linguistic complexity and dialectal variations, cultural context, sarcasm, and figurative language, handling code-switching and mixed language content, real-time processing and scalability, privacy and ethical concerns, and the need for scalable systems [16].
Limited labeled datasets are essential for building accurate emotion detection models, as there is a scarcity of publicly available datasets specifically focused on emotions[17]. Linguistic complexity and dialectal variations are also significant challenges, as Arabic is a complex language with various dialects and regional differences. Cultural context plays a crucial role in influencing emotions, with factors such as sarcasm and figurative language posing significant challenges in emotion detection[18]. This is particularly evident in Arabic, which boasts a rich tradition of poetic expression and metaphors. Emotion detection models need to be capable of handling code-switching and understanding emotions in content that includes multiple languages[19]. Additionally, the real-time processing and scalability of emotion detection systems are key considerations, given the vast amounts of data generated on social media platforms. The scalability issue becomes particularly apparent when dealing with large-scale platforms that have millions of active users. Moreover, the detection of emotions in social media content raises ethical and privacy concerns, underscoring the need for the development of resilient and effective models[18], [20], [21].

4. Feature Selection

In order to enhance the effectiveness and productivity of natural language processing tasks like text classification, sentiment analysis, information retrieval, and machine translation, feature selection plays a crucial role. Arabic language, with its intricate structure and diverse forms, presents specific challenges for feature selection[4], [22].. This section examines various strategies for feature selection tailored specifically

for Arabic text, including hybrid models, information gain-based techniques, statistical approaches, and linguistic-based methods.

4.1 Statistical Approaches for Arabic Text Feature Selection

Statistical methods for Arabic text feature selection, including chi-square, mutual information, and correlation-based feature selection, are examined in this section [23]. Application of the chi-square approach to feature selection in Hadith text classification, with a focus on teaching-type-based categorization. In order to increase classification accuracy, the study shows how well K-nearest neighbor (KNN) and chi-square feature selection work together. This research sheds light on the real-world uses for statistical techniques like chi-square when managing Arabic text data. It aims to enhance the effectiveness of Arabic text mining models by concentrating on the most distinguishing and enlightening characteristics, while also decreasing the dimensionality of the feature space. These methods can identify the most informative features that are highly relevant for sentiment classification, even in the presence of Arabic's rich morphology, dialectal variations, and idiomatic expressions [24][25][26][27]. The statistical feature selection can be performed quickly, often in a single pass over the data, and the selected features can be efficiently integrated into the online sentiment analysis model.

4.2 Information Gain-Based Methods

To choose appropriate features in Arabic text, such as gain ratio and entropy-based feature selection. It investigates how these techniques enhance the efficiency of natural language processing jobs and deal with the particularities of Arabic text[24]. . Dealing with the Complexity of Arabic Language [26]. Tackling the Issue of Limited Data, Grasping Contextual Meanings, and Reducing Complexity [28] [26]. Information Gain-Based methods are generally more efficient in terms of computing compared to other feature selection methods, making them suitable for online or real-time sentiment analysis of Arabic text. The Information Gain calculation can be done quickly, often in a single pass over the data, and the chosen features can be effectively incorporated into the online sentiment analysis model. This enables quick processing and low latency, which are crucial for real-time sentiment analysis applications, as well as offline sentiment analysis of Arabic text [28].

4.3 Linguistic-Based Techniques

This section explores linguistic-based techniques for Arabic text feature selection, such as part-of-speech filtering, semantic role labeling, and syntactic analysis. It examines how linguistic features contribute to the selection of relevant features in Arabic text[29] . Utilizing context-based and category-based relations to measure semantic relatedness between Arabic terms, [30]This approach provides insights into how linguistic features, such as semantic relationships, can be leveraged to enhance the selection of relevant features in Arabic text [31]. The selection of linguistic-based features offers results that are easier to understand and explain, as the features chosen align with our knowledge of the Arabic language . Implementing Linguistic-Based Techniques Methods for online or real-time sentiment analysis of Arabic text can be difficult due to the computational complexity and the requirement for swift processing. Extracting and selecting linguistic features can be a lengthy process, particularly for extensive, streaming data[32]. This approach is well-suited for tasks such as sentiment analysis of archived social media posts, customer reviews, news articles, or other textual data that can be processed in batches.

4.4 Hybrid Models for Arabic Text Feature Selection

it combines various feature selection methods to improve the selection of informative features in Arabic text [21] [33]. There are many hybrid models, such as the hybrid classification model that combines Convolutional Neural Networks (CNN) and Genetic Algorithms [34] , and also a transition from traditional TF-IDF to advanced models like BERT. It presents examples of hybrid models, such as the combination of

statistical and linguistic-based approaches, and highlights their advantages in handling the complexity of Arabic text [35] .

5. Emotion Detection Techniques

Detecting emotions within the Arabic language is an emerging field, but there has been more focus on detecting emotions in English text and developing lexicon resources with new models and techniques [36]. Emotion detection in Arabic is significantly essential as there are a very small number of studies in the literature so far. In the majority of NLP emotion detection studies, it is enough that the methods do not correctly predict and identify seventeen levels of emotion. For that reason, it is also important to specify which emotion with real examples is detected wrong in the content when using emotion detection. This is one of the most important indicators for determining the applicability of the model. The main challenge of sentiment analysis and emotion detection is to develop algorithms that are designed to parse human-defined categories while affecting differences in emotional understanding in expressions. Additionally, contextual information about the problem is one of the crucial factors to be considered[11].

5.1. Lexicon-Based Approaches

Lexicon-Based models are relatively simple and straightforward to implement, as they rely on pre-existing lexical resources rather than requiring extensive training on large datasets. They are commonly used in various NLP applications, Lexicon-based approaches for emotion detection have been thoroughly studied and are often employed with Arabic social media. Generated by collecting Arabic tweets with positive and negative emotions, and calculating sentiment scores for words using pointwise mutual information. Translated the NRC Emotion Lexicon, Bing Liu Lexicon, and MPQA Subjectivity Lexicon into Arabic using Google Translate [37]. Arabic is a rich language with unique grammar and many dialects across different regions[11]. Extracting lexical features from the input text, such as word frequencies, n-grams, parts of speech, or named entities [38].

5.2. Machine Learning Models

Machine learning may be a quickly advancing field that's vital in different applications, counting characteristic dialect handling (NLP), sentiment analysis, and feeling acknowledgment. With the headway of advances just like the Web of Things (IoT) and manufactured insights, the request for solid human acknowledgment frameworks is expanding [39]. In the context of text classification, machine learning algorithms are used to convert text documents into digital representations for tasks such as classifying Arabic text [40] . Techniques such as convolutional neural networks (CNNs) have proven particularly effective for tasks such as English text translation and classification [41] [42]. Researchers have been investigating the effectiveness of different machine learning algorithms, preprocessing techniques, and representation methods to enhance the performance of Arabic document classification[26] [43]. In the realm of sentiment analysis, machine learning serves as the backbone for artificial intelligence and NLP, enabling systems to predict and interpret sentiments accurately[43]. Studies have indicated that combining linguistic features with machine learning algorithms like Support Vector Machines (SVM) and Deep Learning can achieve high performance in tasks such as detecting psychological states[44]. Furthermore, research has focused on the use of lexicon-based and machine learning-based techniques for text representation in sentiment analysis[45] . Deep-learning techniques have outperformed traditional machine-learning techniques in extracting sophisticated features from data. Several recent strategies emphasize the significance of end-to-end learning. These strategies recommend seasoning neural networks so that the model can directly reveal target concepts. Deep learning is a discipline of computer science that seeks to build sophisticated machine learning methods that can automatically learn feature handling and transformation of complex distributions to address complex learning problems [8]. A significant characteristic of DNNs is that they can identify data characteristics adaptive to defects without resuming

handcrafted characteristics. Deep API networks provide state-of-the-art monitoring capabilities for language testing and word predictions. It is rare to use a system that does not use any release resources. Using a distributed algorithm dramatically raises computing costs, making deep learning training quicker and quicker. Different architectural patterns were used, including a DNN (deep neural network), RNN (recurring neural network), LSTM (long short-term memory), CNN (convolutional neural network), and ULMFit. OpenAI's PLA's have received a remarkable agreement in large-scale literary examination analysis competitions on Twitter. However, the fine-grained E-C-STAR data proposed in this crowdsourcing experiment is a critical feature compared to other English-language data. For additional research in the field, generating and sharing a digital corpus of Arabic-language information is indeed demanded [1]. In early work, support vector machine (SVM) and naive Bayes, cosine similarity, and maximum entropy were employed as effective components [46].

The main difference is where they get their information from. Lexicon-centric models use pre-existing dictionaries or lexicons that contain information about words, phrases, and associated details. On the other hand, machine learning-based NLP models learn patterns and connections from large text datasets without relying on a set lexicon. The training methods also vary greatly. Lexicon-centric models don't require significant training on huge datasets because the lexicon acts as a static knowledge base for processing fresh input text. However, machine learning-based NLP models are trained on massive volumes of text data to discover underlying language patterns and correlations. This variation in training methodologies results in differences in flexibility and adaptation. Lexicon-centric models are more rigid since they are limited by the information and coverage available in the pre-compiled lexicon. Updating or expanding such models frequently necessitates manual vocabulary changes. In contrast, machine learning-based NLP models can be more adaptable and flexible since they can learn new patterns and associations from training data and be fine-tuned or retrained for specific domains or tasks. The management of complicated language problems further distinguishes the two techniques. Lexicon-centric models, which rely on a limited set of predefined rules and lexical information, may struggle with subtle linguistic phenomena such as context-dependent meanings, idiomatic expressions, and ambiguous language. Machine learning-based NLP models, particularly those that use advanced approaches such as transfer learning or contextual embeddings, can capture and comprehend the complexity of natural language.

6. Applications of Emotion Detection in Social Media

Emotion detection goes beyond just recognizing emotions in conversations and has a wide range of applications in healthcare, robotics, and social media. In healthcare, it can assist machines and computers in responding more empathetically to distress messages and identifying mental health issues. Emotion detection is increasingly important in social media due to its unstructured nature and vast amount of available data. It is also utilized in recommendation systems to make human-like decisions and can help in identifying social trends using large datasets. Additionally, emotion detection is used to differentiate between trustworthy and untrustworthy bots in online conversations, with trustworthy bots enhancing a company's reputation and customer satisfaction, while untrustworthy bots engage in fraudulent activities. In conclusion, emotion detection is a rapidly expanding field with diverse applications in various domains[3].

The identification of emotions is a connection point between interactions with computers and is employed in making human-like judgments in suggestion systems. For example, a transcript of call records is one way of improving duplicate calls from the same user[9]. The applications of emotion detection in social media are vast and are employed in places where digitized content of emotions is in the picture. The most straightforward one is to use it in decision-making, such as finding a social trend using a large suite of data and mapping emotional states on it.

7. Datasets

Several datasets for various applications found for emotion detection:+

1. [47] The paper discusses the challenges in emotion detection in Arabic text due to limited annotated datasets, presents an Arabic tweet dataset created for this purpose, describes the methodologies used in dataset preparation, and reports the results of preliminary experiments with an overall accuracy of 68.12%.
2. [48] introduces a new dataset called ExaAEC for emotion recognition on Arabic tweets, applying ELMO and LSTM neural networks to achieve an F1-score of 0.65 in emotion detection.
3. [49] focuses on developing a method for threat detection in social media content, particularly images and comments on Instagram, to aid in tracking and monitoring social media posts for threats.
4. [50]Introduces the first Instagram Arabic corpus for cyberbullying detection, designed to detect offensive language in texts, with the SVM classifier outperforming others with F1 scores of 69% for bullying comments and 85% for positive comments. The dataset is publicly available.
5. [51] The paper offers a dataset for detecting comedy in Arabic text by collecting and annotating Arabic humorous tweets in dialects and Modern Standard Arabic, experimenting with seven Arabic Pre-Trained language models to establish a baseline classification system, with CAMeLBERT-DA identified as the best-performing model achieving an F1 score and accuracy of 72.11%.

9. Future Directions and Emerging Trends

Nowadays, emotion detection in social media text has become an active research area. This can be seen in the increasing number of research papers dedicated to it. As developers face several problems related to this topic, it is essential to design an automated process for emotion perception. The currently proposed models, introduced by novel algorithms and tools, have become more accurate and provide better performance for many languages. However, their adaptation for the Arabic language and the analysis of its specific constraints have been limited and lacking. Consequently, our objective was to select and carefully examine emotions in the Arabic language, predefined knowledge, and predefined labeled corpus gathered from the published grounds of emotions in the Arabic language in the recent human-sourced platform.

10. Conclusion and Key Takeaways

Based on this review, we concluded that the main challenges in developing a model for the detection of emotion in Arabic text are that most of the traditional machine learning models are under-performing as compared to deep learning models while distributed representations of words and syntactic dependencies have been largely employed in the deep learning models. An important insight from the results of datasets is that larger datasets are not always producing better results which further conclude the importance of smaller datasets for building and evaluating a model for emotion detection in the Arabic language.

This article provided a comprehensive review of different approaches used in developing models for emotion detection in Arabic textual data. We have covered all the main types of traditional models and deep learning models and presented an evaluation of their individual performance. In addition, we have highlighted the utilization of deep learning architectures in the improvement of the performances of models for different types of deep learning models.

A vast amount of emotional textual data exists in various open platforms, such as social media and discussion forums, where it is labeled with users' emotions. Furthermore, with the emergence of e-commerce, consumer opinions and reviews on different products are becoming an increasingly important source for opinion mining . With that, various emotion-specific datasets are also being widely available for

different languages, including Arabic. As a result, a plethora of techniques, involving deep learning approaches in most cases, have been developed and employed by different research to develop the best model for emotion detection in Arabic social media texts performing better than others.

References:

[1] O. Oueslati, E. Cambria, M. Ben HajHmida, and H. Ounelli, "A review of sentiment analysis research in Arabic language," Futur. Gener. Comput. Syst., vol. 112, pp. 408–430, 2020.
[2] H. J. Aleqabie, M. S. Sfoq, R. A. Albeer, and E. H. Abd, "A Review Of Text Mining Techniques: Trends, and Applications In Various Domains," Iraqi J. Comput. Sci. Math., vol. 5, no. 1, pp. 125–141, 2024.
[3] S. Kusal, S. Patil, J. Choudrie, K. Kotecha, D. Vora, and I. Pappas, "A Review on Text-Based Emotion Detection -- Techniques, Applications, Datasets, and Future Directions," 2022.
[4] F. N. A. Hussein and H. J. Aleqabie, "Cyberbullying Detection on Social Media: A Brief Survey," in 2023 Second International Conference on Advanced Computer Applications (ACA), 2023, pp. 1–6.
[5] B. Abu-Salih et al., "Emotion detection of social data: APIs comparative study," Heliyon, vol. 9, no. 5, 2023.
[6] A. Abaskohi, N. Sabri, and B. Bahrak, "Persian Emotion Detection using ParsBERT and Imbalanced Data Handling Approaches," 2022.
[7] G. Alqahtani and A. Alothaim, "Emotion Analysis of Arabic Tweets: Language Models and Available Resources.," Front. Artif. Intell., vol. 5, p. 843038, 2022.
[8] M. Mhamed, R. Sutcliffe, X. Sun, J. Feng, E. Almekhlafi, and E. A. Retta, "Improving Arabic Sentiment Analysis Using CNN-Based Architectures and Text Preprocessing.," Comput. Intell. Neurosci., vol. 2021, p. 5538791, 2021.
[9] R. Khalaf, M. Makary, I. Technology, and I. Technology, "L EBANON U PRISING : A THOROUGH STUDY OF L EBANESE TWEETS," 2020.
[10] A. Alrumaih, R. Alsabah, H. J. Aleqabie, A. Y. Mjhool, A. Al-Sabbagh, and J. Baldwin, "Analyzing user behavior and sentimental in computer mediated communication," Advances in Intelligent Systems and Computing, vol. 1200 AISC. pp. 287–298, 2021.
[11] K. Darwish et al., "A panoramic survey of natural language processing in the Arab world," Commun. ACM, vol. 64, no. 4, pp. 72–81, 2021.
[12] H. M. Alkhateeb and A. Alnahhas, "Emotion detection in Arabic texts extracted from twitter network by using machine learning techniques," J. Artif. Intell. Mach. Learn. Neural Netw., no. 21, pp. 27–36, Jan. 2022.
[13] N. AlShammari and A. AlMansour, "Aspect-based Sentiment Analysis and Location Detection for Arabic Language Tweets," Appl. Comput. Syst., vol. 27, no. 2, pp. 119–127, 2022.
[14] M. Baali and N. Ghneim, "Emotion analysis of Arabic tweets using deep learning approach," J. Big Data, vol. 6, pp. 1–12, 2019.
[15] N. Hicham, S. Karim, and N. Habbat, "Customer sentiment analysis for Arabic social media using a novel ensemble machine learning approach," Int. J. Electr. Comput. Eng., 2023.
[16] A. Abugharsa, "Sentiment Analysis in Poems in Misurata Sub-dialect," J. Chem. Inf. Model., vol. 53, no. February, p. 2021, 2021.
[17] N. Lim, "Cultural differences in emotion: differences in emotional arousal level between the East and the West," Integr. Med. Res., vol. 5, no. 2, pp. 105–109, 2016.
[18] T. Maqsood, O. Khalid, R. Irfan, S. A. Madani, and S. U. Khan, "Scalability Issues in Online Social Networks," ACM Comput. Surv., vol. 49, no. 2, Sep. 2016.
[19] I. Abu Farha and W. Magdy, "From Arabic Sentiment Analysis to Sarcasm Detection: The

ArSarcasm Dataset," in Proceedings of the 4th Workshop on Open-Source Arabic Corpora and Processing Tools, with a Shared Task on Offensive Language Detection, 2020, pp. 32–39.
[20] M. E. M. Abo, R. G. Raj, A. Qazi, and A. Zakari, "Sentiment Analysis for Arabic in Social Media Network: A Systematic Mapping Study," 2019.
[21] A. Katirai, "Ethical considerations in emotion recognition technologies: a review of the literature," AI Ethics, pp. 1–22, 2023.
[22] D. J. Mohammed and H. J. Aleqabie, "Context-Based Visual Sentiment Analysis for Social Media Data," in 2022 International Conference on Artificial Intelligence of Things (ICAIoT), 2022, pp. 1–5.
[23] N. Yang and Y. Zhang, "Railway Fault Text Clustering Method Using an Improved Dirichlet Multinomial Mixture Model," Math. Probl. Eng., 2022.
[24] S. Bahassine, A. Madani, M. Al-Sarem, and M. Kissi, "Feature Selection Using an Improved Chi-Square for Arabic Text Classification," J. King Saud Univ. - Comput. Inf. Sci., 2020.
[25] G. Z. Nabiilah, S. A. Faraby, and M. D. Purbolaksono, "Classification of Hadith Topic of Indonesian Translation Using K-Nearest Neighbor and Chi-Square," Int. J. Inf. Commun. Technol., 2021.
[26] A. Y. Muaad et al., "Arabic Document Classification: Performance Investigation of Preprocessing and Representation Techniques," Math. Probl. Eng., 2022.
[27] Y. Zhu, Z. Lü, X. Chen, Y. Li, and J. Wang, "Identification of Cashmere and Wool Based on LBP and GLCM Texture Feature Selection," J. Eng. Fiber. Fabr., 2023.
[28] F. A. Abdulghani and N. A. Abdullah, "A Survey on Arabic Text Classification Using Deep and Machine Learning Algorithms," Iraqi J. Sci., 2022.
[29] W. Etaiwi and A. Awajan, "SemG-TS: Abstractive Arabic Text Summarization Using Semantic Graph Embedding," Mathematics, 2022.
[30] "Exploiting Wikipedia to Measure the Semantic Relatedness Between Arabic Terms," J. Eng. Res. Technol., 2022.
[31] W. Etaiwi and A. Awajan, "Graph-Based Arabic Text Semantic Representation," Inf. Process. Manag., 2020.
[32] J. Gong et al., "Hierarchical Graph Transformer-Based Deep Learning Model for Large-Scale Multi-Label Text Classification," Ieee Access, 2020.
[33] ه. محمد الخطيب, "Emotion detection in Arabic texts extracted from twitter network by using machine learning techniques," مجلة المعهد العالي للدراسات النوعية, vol. 3, no. 2, pp. 527–541.
[34] D. Alsaleh and S. L. Marie-Sainte, "Arabic Text Classification Using Convolutional Neural Network and Genetic Algorithms," Ieee Access, 2021.
[35] N. Al-Twairesh, "The Evolution of Language Models Applied to Emotion Analysis of Arabic Tweets," Information, 2021.
[36] S. K. Bharti et al., "Text-Based Emotion Recognition Using Deep Learning Approach.," Comput. Intell. Neurosci., vol. 2022, p. 2645381, 2022.
[37] S. M. Mohammad, M. Salameh, and S. Kiritchenko, "Sentiment lexicons for Arabic social media," Proc. 10th Int. Conf. Lang. Resour. Eval. Lr. 2016, pp. 33–37, 2016.
[38] L. Almuqren, R. Hodgson, and A. I. Cristea, "Arabic Text Sentiment Analysis: Reinforcing Human-Performed Surveys with Wider Topic Analysis."
[39] S. Pal, S. C. Mukhopadhyay, and N. K. Suryadevara, "Development and Progress in Sensors and Technologies for Human Emotion Recognition," Sensors, 2021.
[40] F. A. Abdulghani and N. A. Z. Abdullah, "A Survey on Arabic Text Classification Using Deep and Machine Learning Algorithms," Iraqi J. Sci., vol. 63, no. 1, pp. 409–419, 2022.
[41] M. Alhawarat and A. O. Aseeri, "A Superior Arabic Text Categorization Deep Model (SATCDM)," Ieee Access, 2020.
[42] D. Bouchiha, A. Bouziane, and N. Doumi, "Machine Learning for Arabic Text Classification: A

Comparative Study," Malaysian J. Sci. Adv. Technol., 2022.
[43] Z. Anwar, R. Jahangir, M. A. Nauman, R. Alroobaea, S. M. Alzahrani, and I. Ali, "Context-Based Emotion Predictor: A Decision- Making Framework for Mobile Data," Mob. Inf. Syst., vol. 2022, pp. 1–12.
[44] X. Du and Y. Sun, "Linguistic Features and Psychological States: A Machine-Learning Based Approach," Front. Psychol., 2022.
[45] J. Mutinda, "Sentiment Analysis on Text Reviews Using Lexicon Selected-Bert Embedding (LeBERT) Model With Convolutional Neural Network," 2022.
[46] N. Ashraf, L. Khan, S. Butt, H.-T. Chang, G. Sidorov, and A. Gelbukh, "Multi-label emotion classification of Urdu tweets.," PeerJ. Comput. Sci., vol. 8, p. e896, 2022.
[47] A. Al-Khatib and S. R. El-Beltagy, "Emotional tone detection in Arabic tweets," in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2018, vol. 10762 LNCS, pp. 105–114.
[48] S. Sarbazi-Azad, A. Akbari, and M. Khazeni, "ExaAEC: A New Multi-label Emotion Classification Corpus in Arabic Tweets."
[49] S. A. AlAjlan and A. K. J. Saudagar, "Machine learning approach for threat detection on social media posts containing Arabic text," Evol. Intell., vol. 14, no. 2, pp. 811–822.
[50] R. ALBayari and S. Abdallah, "Instagram-Based Benchmark Dataset for Cyberbullying Detection in Arabic Text," Data, vol. 7, no. 7, p. 83.
[51] H. Alkhalifa, F. Alzahrani, H. Qawara, R. AlRowais, S. Alowa, and L. Aldhubayi, "A Dataset for Detecting Humor in Arabic Text," Int. Conf. Nat. Lang. Speech Process.