**Indonesian Journal of Electrical Engineering and Computer Science**
Vol. 40, No. 3, December 2025, pp. 1500~1516
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v40.i3.pp1500-1516 

# Sentiment analysis in Arabic and dialects: a review utilizing a corpus-based approach

**Abbas Hussein Ali[1,2], Necaattin Barişçi[3]**
[1]Department Of Computer Science, Graduate School of Informatics, Gazi University, Ankara, Turkey
[2]Department of Computer Science, Faculty of Computer Science and Information Technology, Kirkuk University, Kirkuk, Iraq
[3]Department of Computer Engineering, Faculty of Technology, Gazi University, Ankara, Turkey

**Article Info**

***Article history:***

Received Mar 14, 2025
Revised Oct 15, 2025
Accepted Nov 16, 2025

***Keywords:***

Arabic dialects
Deep learning
Machine learning
Natural language processing
Sentiment analysis

**ABSTRACT**

Arabic is one of the most morphologically complex languages, and its numerous dialects render identifying sentiment in digital communication a challenging task. In this study, we conduct a systematic literature review (SLR) to investigate the sentiment analysis (SA) techniques used on modern standard Arabic (MSA) and several Arabic dialects (AD) between 2020 and 2024. A corpus-based analysis of 71 articles indicated that machine learning (ML) and deep learning (DL) algorithms were the dominant methods used. Overall, the most frequently studied dialects are those from Saudi Arabia, Morocco, and to a lesser extent, Algeria, among various algorithms used for text classification, including support vector machines (SVM) and convolutional neural networks (CNN). These techniques emerged as some of the most effective strategies employed for sentiment classification. While new contemporary word embeddings, such as Word2Vec, are gaining traction in the field, traditional feature extraction methods, like term frequency-inverse document frequency (TF-IDF), continue to outperform them. The study highlights the importance of additional labeled datasets and tailored models in navigating the linguistically rich world of AD. Additionally, the results highlight the need for dialect-specific adaptations to improve SA outcomes, and further investigation is needed by leveraging advanced DL methodologies, as well as improved data resources, to address these issues.



***Corresponding Author:***

Abbas Hussein Ali
Department of Computer Science, Faculty of Computer Science and Information Technology
Kirkuk University
Sayade district, Kirkuk, Iraq
Email: abbasalsaka@uokirkuk.edu.iq

## 1. INTRODUCTION

As computers continue to advance, their ability to process and compute becomes the cornerstone of modern artificial intelligence [1]. Natural language processing (NLP) enables machines to understand human language, empowering applications such as translation, summarization, and text classification for effective data organization and analysis [2]. SA is a vital and active research domain in NLP, essential for interpreting human emotions and attitudes expressed in digital text [3]. It offers significant insights into subjective content, enhancing our understanding of human emotions and opinions [4]. In recent instances, researchers have shown an increasing interest in Arabic natural language processing (ANLP), which incorporates various packages, including machine question answering (QA), automatic machine translation (ATM), and sentiment analysis, among others. However, despite the wealth of online Arabic statistics, the sphere of ANLP, especially in sentiment analysis, faces challenges stemming from the limited availability of resources compared to English [5]. The morphological complexity of the Arabic language is widely recognized, as it holds the distinction of being one of the official languages of the United Nations and is spoken by a vast global population. In addition, it is currently placed as the fourth most widely used language on the internet [6]. With a population of 422 million, Arabic serves as the primary language. Arabic is written from right to left. The Arabic language consists of three principal variations: Classical Arabic (CA), that is used in the Holy Qur'an; MSA and DA [7].

The usage of DA in speech is informal and has significant variations across different regions and nations. MSA is utilized exclusively in various scenarios, such as news broadcasts, written materials, formal addresses, film subtitles, and religious practices. It naturally functions as the shared language among Arabs, often used in conversation to enhance universal understanding. On the other hand, dialectal Arabic refers to the colloquial language used in informal contexts, including business, social circles, and family environments. The Arab world's community reveals an attraction for deploying their own dialects in everyday conversations, resulting in a growing number of ADs in social media, radio programmers, and television shows [8].

In contrast, it has been found that a variety of AD are spoken in different countries, despite Arabic being widely used as the primary language in the majority of Arab countries. The absence of standardized written forms in these dialects, combined with differences between Arab countries, presents challenges for automated processing. The diversity observed across ADs throughout the Arab world may be mainly attributed to slight variations in syntax and vocabulary, which subsequently provide rise to various pronunciations and writing rules [9]. The Arabic language received the attention of academics as a result of the escalating usage rate of the language on the internet; therefore, an enormous amount of daily online information is generated in the Arabic language. As a result, there is a growing demand for reliable technology and efficient methodologies to process texts written in either MSA or AD [10].

The research paper is structured into several sections: the second section describes the research methodology, the third section introduces background concepts, the fourth section discusses related works and previous findings, the fifth section focuses on results and provides detailed analysis, and the sixth section discusses the results and the final section presents the conclusions drawn from the study.

## 2. METHOD

Sentiment analysis, also known as opinion mining, is a computational procedure that involves identifying and classifying opinions presented in written material into positive, negative, or neutral categories. The process involves examining written content to determine the sentiment or emotion expressed by the writers. The application of this analysis is prevalent across diverse applications, including but not limited to social media monitoring, market research, and customer feedback analysis [11].

This SLR investigates ML and DL techniques used for SA between 2020 and 2024, with emphasis on Arabic dialect. Figure 1 illustrates the scope of the study. To ensure a scientific and comprehensive presentation, this study utilized the preferred reporting items for systematic reviews and meta-analyses (PRISMA) standard, as illustrated in Figure 1. In the identification stage, 264 articles were initially identified. From the initial 264 articles, 245 were selected for screening. In the eligibility stage, the remaining articles were examined more thoroughly to assess their relevance to the research scope, and only 123 were deemed eligible for further consideration. Finally, only 71 studies were included in the SLR based on their quality and relevance.

To conduct the SLR on arabic sentiment analysis (ASA), we first compiled a list of relevant terms and keywords used in the ASA. We then utilized these keywords to initiate a search for relevant manuscripts using the most popular database. Our search criteria included manuscripts published in English between 2020 and 2024 that contained at least one of the following search strings in their titles: 'Arabic Dialects', 'Sentiment Analysis', 'Colloquial Arabic'. After conducting the search process, we obtained 264 articles. Subsequently, we screened their abstracts, removed duplicates and irrelevant records, resulting in a final selection of 71 manuscripts for the analysis.

There were 71 papers in all chosen from several databases. ScienceDirect yielded the most papers 23, followed by Springer with 10 papers. IEEE Xplore and the ACM digital library each contributed 7 and 8 papers, respectively. ArXiv and Clarivate also each provided four papers. While the MDPI Journal counted five papers, other databases, including ACL Anthology, Academia, and Scopus, each provided three papers. Hindawan also provided one paper. This distribution emphasizes the variety of materials used in the research.

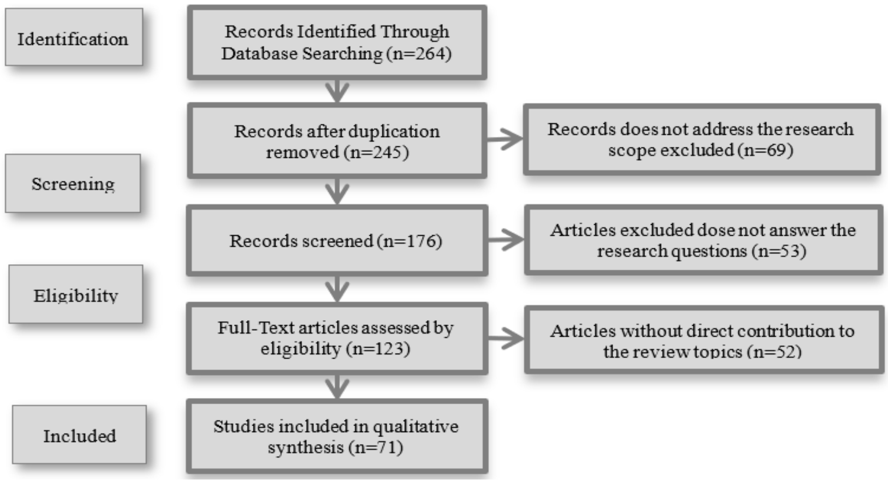


Figure 1. Information flow in systematic review stages

Research questions: The objective of this study is to provide answers to several questions that can assist researchers in selecting the most suitable approach for ASA using ML and DL techniques.
**RQ1**/Which Arabic language variety represents the most significant proportion in the studies?
**RQ2**/Which ML and DL algorithms are predominantly utilized in studies of AD?
**RQ3**/What are the primary input characteristics employed in studies of AD?
**RQ4**/What types of data sources are predominantly utilized in studies aimed at identifying AD?

## 3. BACKGROUND

### 3.1. Arabic language

Millions of people speak the Arabic language, including its various dialects and MSA. The geographical scope of the Arabic language is significant, spanning from the Atlantic Ocean to the Persian Gulf. Arabic is the primary language for many people and a secondary or shared language in Muslim regions worldwide. Structurally, the total twenty-eight letters of the Arabic alphabet include three long vowels and twenty-five consonants. Arabic script is written from right to left, and the shape of Arabic letters changes depending on where they are in a word, which makes it different from English [12]. MSA is utilized in educational, news, and governmental platforms, while the dialects are used in daily communication and digital interaction [13].

Arabic, being a Semitic language with a variety of dialects that depend on specific regional areas, presents unique problems for NLP. Alongside CA and MSA as official variations, each of the 22 Arab states possesses its own dialect, leading to a diversity of dialects that reflect regional differences. The phonology, vocabulary, and syntax of standard dialects differ significantly from those of CA and MSA. Differences can give rise to socially controversial discussions. Nevertheless, Arabic culture and language are rich and diverse [6]. The Arab region plays a significant role in the context of global politics and the worldwide economy. Therefore, it receives a lot of attention from international institutions that want to understand opinions related to various subjects, such as oil pricing, the stock market, and foreign policy issues [14]. NLP encounters special problems when dealing with the Arabic language, mainly due to its complex morphology, numerous dialects, and right-to-left writing. Due to the large number of Arabic dialects that lack standardized orthography, there are few resources available, and regional diversity strategies are required to increase comprehension. Strong ANLP systems are restricted by grammatical ambiguity, dialectal variations, and the lack of available language resources [15].

### 3.2. Sentiment analysis

The growth of available information has led to the development of SA, which has evolved as an organized structure for recognizing and categorizing sentiments. Recognition of patterns is the primary objective of SA, classifying statements as positive, negative, or neutral. However, the scope of this attempt may expand beyond ordinary polarity to cover specific sentiments and emotions, including fear, joy, and anger [16]. Recent advancements in ML have improved the accuracy and effectiveness of SA and recommendation systems [17]. SA employs several methodologies for text analysis. In general, SA functions at three primary levels, namely the document-level, sentence-level, and aspect-based order. At the document level, the study assumes that the entire document expresses a single opinion. While at the sentence level, each sentence is treated as an individual opinion, and the overall sentiment of the document depends on the sentiment of its sentences. Additionally, at the aspect-based level, a single sentence may include multiple aspects, each with its own distinct opinion [18]. As stated by [19], there is a modern trend where people share their opinions, ideas, and reviews on various topics related to products or services across different internet platforms, including websites, blogs, and social media. Researchers in the field of NLP have demonstrated significant interest in the extensive content created by individuals [20]. According to the research [21], internet platforms have become important sources of valued opinions. These platforms contribute to the needs of companies that are interested in obtaining public evaluations on various products or services. Dialectal languages are commonly utilized by individuals in daily interactions throughout the Arab world. Moreover, they frequently utilize dialectal forms of language to express their thoughts and ideas on the internet. This results in the production of a considerable quantity of dialectal Arabic texts, which presents a difficulty for ANLP [10].

### 3.3. Feature extraction techniques

Data preprocessing is crucial in data analysis, as it transforms raw data into a structured, error-free format. Key steps include data cleaning, error correction, and data transformation techniques like scaling and feature selection for enhanced interpretability and quality [22]. The transformation of raw data into numerical attributes that can be managed while preserving the integrity of the original data is referred to as feature extraction. In text analysis, it is necessary to convert the text input into numerical form because ML algorithms work more effectively with numerical input than with text input. The goal of the feature extraction approach is to identify the most significant aspects of a text so that relevant information may be extracted from it [23]. During the feature extraction process, features that are determined to be irrelevant or unnecessary will be removed. Using feature extraction to prepare data for a learning algorithm might improve the accuracy of the system while also saving the amount of time necessary for it to learn [24]. Word embedding is a feature extraction approach that involves representing words as numeric vectors. These vectors represent words that have similar meanings and scenarios. In contemporary times, word embeddings are frequently employed in various NLP tasks. In DL approaches, they often function as input layers [25]. The Word2Vec model, created by Google, is currently recognized as the most widely used word embedding model. Word2Vec incorporates two learning models, specifically the continuous bag-of-words (CBOW) and skip-gram (SG) models. In the CBOW model, contextual words inside the surrounding context are employed to provide predictions for a target word. In contrast, the SG technique relies on a keyword to predict the surrounding context words [26].

Stanford researchers developed global vectors for word representation (GloVe) for unsupervised learning. It merges global word co-occurrence matrices from a corpus to construct word embeddings using statistics to uncover word associations. Each value in co-occurrence matrices represents a frequently co-occurring word pair, unlike occurrence matrices. The approach computes text co-occurrence probabilities to reveal word correlations. Another technique used for word representation is TF-IDF, which assesses the significance of a word based on its relevance to a document within a group of documents. This method involves multiplying the number of times a word appears in a document (TF) by the word's IDF across a group of documents. These scores can be used to highlight important words that stand out in a document. TF-IDF offers a method for assigning each word in a text a numerical value indicating its significance to the document as a whole [27].

## 4. RELATED WORKS

The study examined several parameters for each article, including information on the publisher, year of publication, and sources of data. In addition, the size of the datasets used was listed, along with the source from which they were obtained, indicating different types of dialects. Additionally, the study examined the feature extraction technique employed for sentiment analysis, differentiating between ML and DL approaches. Both ML and DL classifiers were carefully investigated, with a primary focus on performance indicators such as accuracy, to determine the most effective classifiers, as demonstrated in Table 1 in APPENDIX.

## 5. RESULTS

The dataset that was considered for analysis comprises 71 articles on ASA and was released between 2020 and 2024. According to the four original research questions that were presented at the beginning of this review, the subsequent section will be structured in accordance with those questions.

**RQ1**/ Which Arabic language variety represents the most significant proportion in the studies?

The study's findings indicate that the Saudi, Moroccan, and Algerian dialects demonstrate the highest proportions, with each of these languages representing 21%, 18%, and 17%, respectively, of the overall total. MSA studies and multi-dialect studies were analyzed as separate categories to ensure conceptual clarity; these focus areas each represented 9% of the overall studies. Various percentages, ranging from 10% to 4%, are associated with other dialects, as shown in Figure 2.

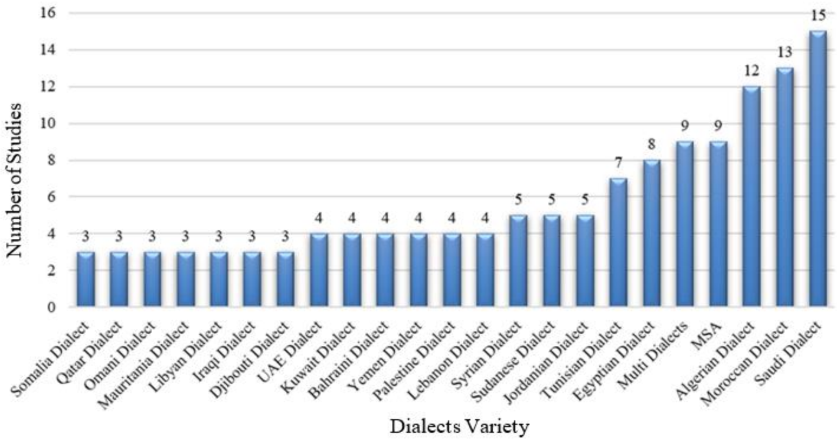


Figure 2. Distribution of selected papers per dialect

**RQ2**/Which ML and DL algorithms are predominantly utilized in studies of ADs?

The study demonstrates a wide range of ML techniques, some of which are seen just once. These algorithms include extra trees, Naive Bayes (NB) classifier, fuzzy C-means clustering, and others. Each of these algorithms accounts for 2% of the dataset. Adaptive boosting, K-means clustering, and stochastic gradient descent (SGD) Classifier algorithms are mentioned twice, accounting for 4% of the total. Extreme gradient boosting was mentioned four times, representing 7% of the total. Multinomial NB accounts for 11% of the total, with a frequency of six occurrences. Support vector machines (SVM) is the most frequently seen model, with 42 occurrences, accounting for 76% of the dataset, as demonstrated in Figure 3. The prevalence of SVM is likely due to its inherent effectiveness with high-dimensional, sparse data, which is characteristic of Arabic text, making it a robust traditional choice for Arabic text classification [28]. The present research highlights the diverse application of ML methodologies, with some algorithms exhibiting a greater prevalence in the existing body of literature compared to others, as demonstrated in Figure 3. Additionally, Figure 4 provides an overview of how ML algorithms are being applied to ASA from 2020 to 2024, highlighting notable trends. Traditional classifiers, such as SVM and NB, are still widely used.

SVM is used 4 to 5 times per year, while NB peaked in 2020 with five uses before stabilizing at 3 uses in 2023. LR is continuously used 3 times per year, serving as a reliable benchmark. Clustering methods, such as random forest (RF), are used 2 to 3 times per year, while others, such as AdaBoost and XGBoost, are used less frequently. The emergence of LightGBM in 2024 indicates increased interest in efficient and scalable clustering techniques. Decision trees have been used in a moderate two applications per year since 2021, often eclipsed by more advanced methods such as RF. Models using gradient descent techniques, such as SGD and NuSVM, experienced a peak in 2020 but declined in 2023.

Moreover, DL algorithms present a diverse range of models that exhibit varying degrees of predominance in ASA. Several algorithms, such as bidirectional long short-term memory (BiLSTM), AlgBERT, Hybrid BiGRU-BiLSTM, multi-layer perceptron (MLP), multilingual adaptable robust augmented bidirectional encoder representations from transformers (MARABERT), enhanced long short-term memory (ELSTM), multiplicative long short-term memory (mLSTM), deep neural networks (DNN), bidirectional gated recurrent unit (BiGRU), and CNN-BiLSTM, each contribute 2% to the overall research. The included articles contain two instances of MLP, GRU, and BiGRU, each accounting for 4% of the total. The CNN-LSTM model appears four times, making up 7% of the research. A recurrent neural network (RNN) is observed in six studies, accounting for 11% of the dataset. AraBERT appears seven times, contributing to 13% of the studies. The bidirectional encoder representations from the transformers (BERT) method have been identified as a prominent choice, since it has been referenced eight times in several academic works, accounting for 15% of the dataset. The RNN model is seen six times, accounting for 11% of the dataset. The BiLSTM model has occurred 11 times, accounting for 20% of the total dataset, while the LSTM model has been observed 14 times, making up 25%. Finally, CNN exhibits the highest frequency, with a total of 16 occurrences, accounting for 29% of the academic studies, as illustrated in Figure 5.

The prominence of CNNs stems from their ability to capture effective local features and sequential patterns from text embeddings, which is important for identifying specific slang, negation, or sentiment phrases within the variable morphology of Arabic dialects [6]. The present research highlights the heterogeneous distribution of ML and DL algorithms across the existing body of literature, whereby several models exhibit more prominence than others. Moreover, Figure 6 illustrates the application of DL algorithms for Arabic dialect sentiment analysis, spanning the years 2020 to 2024. The trends show a drift toward transformer-based models, especially AraBERT and BERT, which became very popular during the later years due to their capability of handling the contextual and morphological complexity of text [29]. CNN and LSTM maintain a constant level of usage from 2 in 2020 to 4 in 2023, with a slight decrease in 2024. AraBERT and BERT just show a growing interest in pre-trained and contextualized transformers. The graph shows the supremacy of transformer variants for NLP tasks. In this evolution, it was shown how the reliance on advanced DL models has grown to solve the challenges posed by the AD for SA.

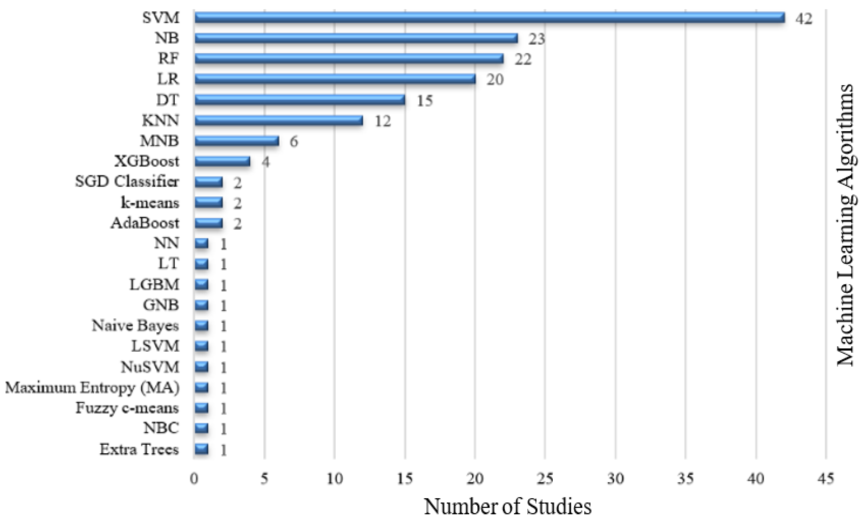


Figure 3. Machine learning algorithms usage by selected papers

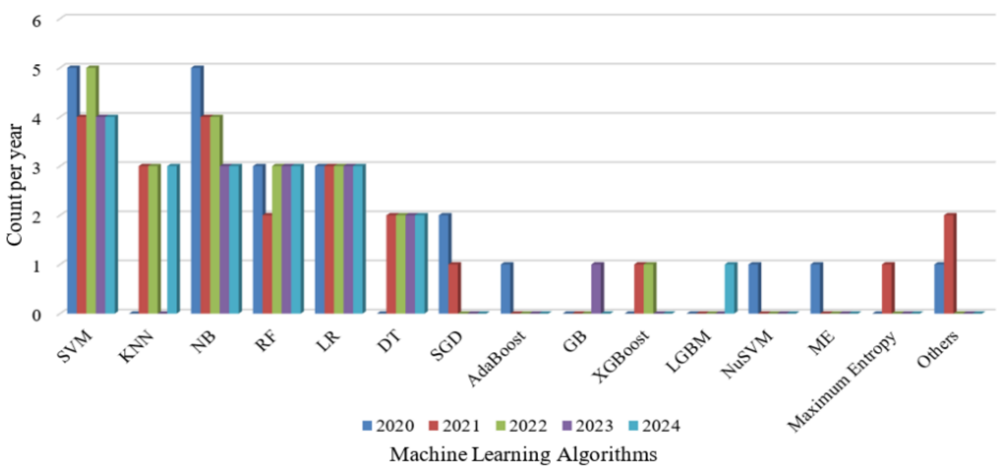


Figure 4. Machine learning algorithms usage per years

**RQ3**/What are the primary input characteristics employed in studies of AD?

The analysis of included studies reveals a diverse range of techniques used across the volume of research. The Terms-document matrix (TDM) was used by the studies three times, accounting for 5% of the total. AraBert and AraVec both have dual appearances, with each model accounting for 4% of the research. GloVe also appears three times, representing a proportion of 5% of the total occurrences. The SG technique was observed four times, accounting for 7% of the total. FastText is used in six research studies, making up 11% of the total dataset, while n-grams were used in seven studies, representing 13%. The CBOW model is referenced in eight studies, constituting 15% of the total occurrences. In contrast, the Word2Vec method emerges as a prominent choice, appearing 14 times and representing 25% of the sample. The bag of words (BOW) approach is utilized even more frequently, with 20 instances, thereby accounting for 36% of the studies analyzed. However, the most widely adopted technique was TD-IDF; it was used in 29 research, representing 53% of all occurrences. This investigation highlights the wide range of text representation approaches used in academic conversations, with certain strategies being more common and important than others, as shown in Figure 7.

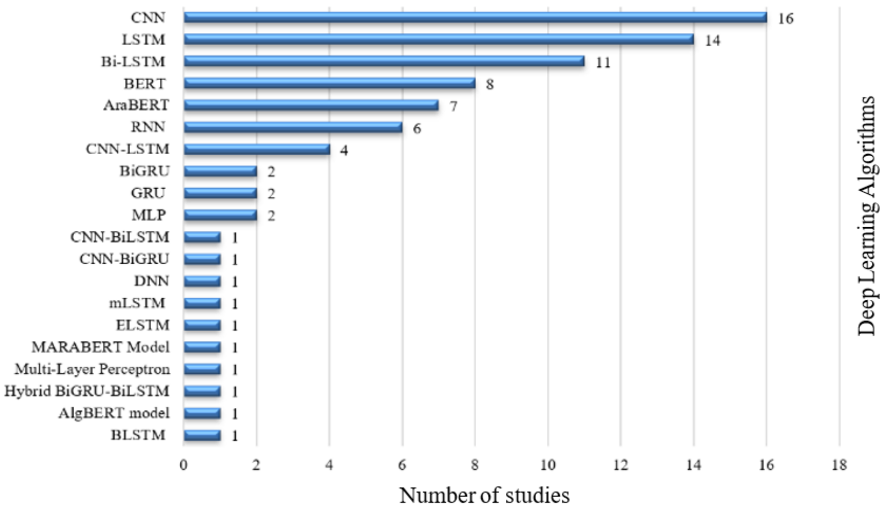


Figure 5. Deep learning algorithms usage by selected papers

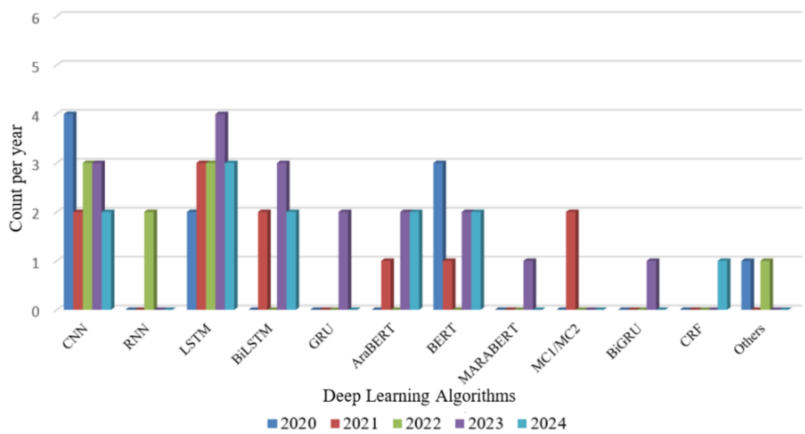


Figure 6. Deep learning algorithms usage per years

**RQ4**/What types of data sources are predominantly utilized in studies aimed at identifying AD?

The examination of the provided data sources reveals a varied range in terms of counts and percentages. With 33 occurrences, Twitter emerges as the most popular source, comprising 46% of the dataset. On the other hand, the Google Play Store was used as a data source three times, representing 4% of the total. While online reviews contributed nine instances, making up 13% of the reviewed research. Likewise, Facebook is mentioned six times, constituting 8% of the dataset. The mix social media category, which comprises the integration of multiple platforms, is particularly notable for its 19 instances, accounting for 27% of the dataset. YouTube and TripAdvisor comments are comparatively infrequent, each manifesting merely once and contributing 1% to the overall count as demonstrated in Figure 8. This analysis highlights the diverse range of sources that have contributed to the dataset. Twitter and mix social media, in particular, have made substantial contributions, while additional data from other sources has enhanced the overall diversity and scope of the dataset.

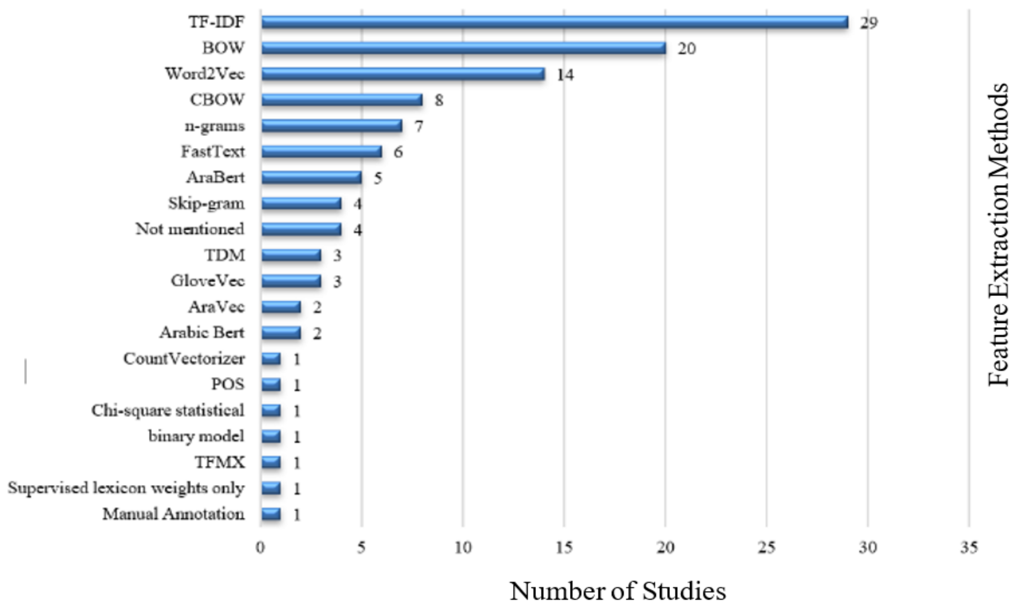


Figure 7. Feature extraction methods used by papers

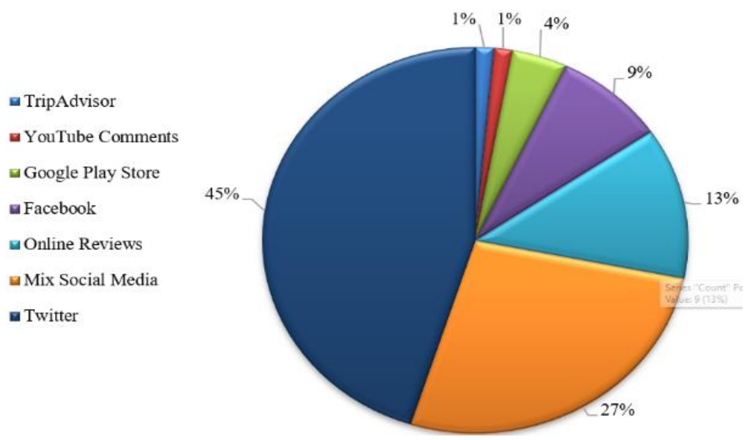


Figure 8. Dataset source among papers

## 6. DISCUSSION

The systematic review of ASA highlights critical research priorities and methodological trends that define the current state of the field. The review found that the concentration of research in the Saudi, Moroccan, and Algerian dialects, with each occurring 15, 13, and 12 times respectively, reveals a significant resource gap for the remainder of the Arab world. In contrast, dialects such as Somali, Qatari, and Omani were examined in only three instances each. This disparity suggests that current ASA models may lack generalization and robustness to Arabic dialect text, emphasizing the need for publicly available datasets to advance the field. The main reason for this reduction is still the complex morphology of the language and the wide diversity of dialects. The increasing prevalence of social media in Arabic-speaking regions has garnered significant academic interest in the domain of ASA. However, several significant research shortcomings remain hindering its progress. A primary challenge is the scarcity of annotated datasets, particularly for underrepresented dialects. The methodological findings demonstrate a clear progression in the field, with the choice of feature representation intrinsically linked to the choice of classifier. The enduring effectiveness of SVM as the most utilized ML algorithm is intrinsically linked to the high frequency of TF-IDF and BOW feature extraction. The reviewed papers consistently show that SVM excels in classifying data within the high-dimensional, sparse feature space that TF-IDF generates when processing morphologically complex Arabic text [24].

In contrast, the high usage of DL algorithms like CNN, LSTM, and Bi-LSTM signals a necessary progression toward more sophisticated, automated feature learning. These models require a contemporary word embedding, which provides vector representations that can inherently manage the ambiguity and morphological variation of ADs. The steady use of CNNs, in particular, confirms their effectiveness in automatically capturing local sentiment features and sequential patterns that are crucial in informal, context-dependent communication, often resulting in superior performance over traditional methods in dialectal text classification [30].

This comparative trend shows that researchers are now preferring models that can autonomously learn semantic and syntactic nuances, rather than relying solely on manually engineered features. Even though advanced DL models like CNNs and LSTMs show great promise, many still depend on traditional ML methods like SVMs and NB. Exploring hybrid models that combine traditional and DL methods could lead to better performance. Research on pre-trained language models, such as AraBERT and MARABERT, mainly relies on traditional techniques like TF-IDF and BOW for feature extraction. Using these advanced models will help ASA become more robust and accurate. Over-reliance on Twitter as the main data source limits the variety of datasets even more, since other platforms, including Facebook, Google Play Store, and YouTube, are still underused. Including these platforms in expanding data collection will help to give a more complete awareness of sentiment in different contexts. Furthermore, complicating model comparison and benchmarking are the lack of dialect-specific sentiment lexicons and consistent evaluation measures. Establishing shared responsibilities and developing these materials would encourage creativity and guarantee comparability between researchers. By filling in these gaps, future studies can greatly advance the field and provide robust, inclusive solutions for ASA.

This study, while aiming to be comprehensive, is subject to methodological limitations. The strict timeframe of 2020 to 2024 may have excluded older, yet foundational, papers that established initial benchmarks and core methodologies in the field. Furthermore, the reliance on standard academic databases introduces a potential publication bias, meaning that studies reporting positive or state-of-the-art results are likely over-represented. Future reviews should strive to incorporate data from regional proceedings to offer a more comprehensive perspective of the field.

## 7. CONCLUSION

This study investigates research dealing with the SA of the Arabic language between 2020 and 2024, including both dialects and Modern Standard Arabic. We covered a comprehensive assessment of the methodologies followed and outcomes acquired for each research. The study reveals that previous research mostly concentrates on Saudi, Moroccan, and Algerian dialects, highlighting a visible absence of studies for other Arabic dialects. This disparity suggests that current ASA models may lack generalization and robustness across the full spectrum of Arabic dialects, emphasizing the need for publicly available datasets.

The research results revealed that SVM, NB, and RF emerged as the primary ML methodologies, emphasizing their effectiveness in addressing the complexities of the Arabic language. The analysis of feature extraction methodologies revealed a notable inclination towards traditional techniques, specifically TF-IDF and BOW. This preference underscores a continued reliance on established methods. Notably, DL models such as CNN, LSTM, and Bi-LSTM have increasingly contributed to include research. Recent studies have increasingly highlighted the significance of these models, as they excel at identifying intricate patterns within Arabic text data. Their ability to understand and analyze language nuances. As a dataset source, Twitter emerged as the primary data source for SA, accounting for nearly half of the dataset. This highlights its importance in capturing public opinions in Arabic-speaking regions. Whereas contributions from online reviews and Facebook were comparatively lower.

Moving forward, it is paramount to address the critical challenge of data scarcity and model generalization. Future research in SA within the context of ADs should prioritize the pressing need for the development of annotated datasets for underrepresented dialects to achieve a more equitable distribution of data and enhance overall model performance. The adoption of advanced DL techniques and hybrid models is also advisable to strengthen the accuracy and resilience of ASA.

**FUNDING INFORMATION**

The authors state no funding is involved.

## AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

| Name of Author | C | M | So | Va | Fo | I | R | D | O | E | Vi | Su | P | Fu |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Abbas Hussein Ali | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ | | | |
| Necaattin Barişçi | ✓ | | | ✓ | ✓ | | | | | ✓ | | ✓ | ✓ | |

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

## CONFLICT OF INTEREST STATEMENT

The authors state there is no conflict of interest.

## DATA AVAILABILITY

Data availability does not apply to this paper as no new data were created or analyzed in this study.

## REFERENCES

[1] M. Alisawi *et al.*, "Cyber security after COVID 19: a review," in *AIP Conference Proceedings*, 2023, vol. 2839, no. 1, p. 040014, doi: 10.1063/5.0167890.

[2] S. N. Mohammed and S. Gündüç, "GuideWalk: a novel graph-based word embedding for enhanced text classification," *arXiv preprint arXiv:2404.18942*, Sep. 2024, [Online]. Available: http://arxiv.org/abs/2404.18942.

[3] A. Ahmed Bilal, O. Ayhan Erdem, and S. Toklu, "Children's sentiment analysis from texts by using weight updated tuned with random forest classification," *IEEE Access*, vol. 12, pp. 70089–70104, 2024, doi: 10.1109/ACCESS.2024.3400992.

[4] M. Baker, K. Jihad, and Y. Taher, "Prediction of people sentiments on twitter using machine learning classifiers during russian aggression in Ukraine," *Jordanian Journal of Computers and Information Technology*, vol. 9, no. 3, p. 1, 2023, doi: 10.5455/jjcit.71-1676205770.

[5] A. Al-Owisheq, S. Al-Humoud, N. Al-Twairesh, and T. Al-Buhairi, "Arabic sentiment analysis resources: a survey," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9742, 2016, pp. 267–278.

[6] A. B. Nassif, A. Elnagar, I. Shahin, and S. Henno, "Deep learning for Arabic subjective sentiment analysis: Challenges and research opportunities," *Applied Soft Computing*, vol. 98, p. 106836, Jan. 2021, doi: 10.1016/j.asoc.2020.106836.

[7] I. Al-huri, "Arabic language : historic and sociolinguistic characteristics," *English Literature and Language Review*, vol. 1, no. 4, pp. 28–36, 2015, [Online]. Available: http://arpgweb.com/?ic=journal&journal=9&info=aims.

[8] A. Elnagar, S. M. Yagi, A. B. Nassif, I. Shahin, and S. A. Salloum, "Systematic literature review of dialectal Arabic: identification and detection," *IEEE Access*, vol. 9, pp. 31010–31042, 2021, doi: 10.1109/ACCESS.2021.3059504.

[9] Y. Matrane, F. Benabbou, and N. Sael, "A systematic literature review of Arabic dialect sentiment analysis," *Journal of King Saud University - Computer and Information Sciences*, vol. 35, no. 6, p. 101570, Jun. 2023, doi: 10.1016/j.jksuci.2023.101570.

[10] A. A. Al Shamsi and S. Abdallah, "Text mining techniques for sentiment analysis of Arabic dialects: literature review," *Advances in Science, Technology and Engineering Systems Journal*, vol. 6, no. 1, pp. 1012–1023, 2021, doi: 10.25046/aj0601112.

[11] M. Wankhade, A. C. S. Rao, and C. Kulkarni, "A survey on sentiment analysis methods, applications, and challenges," *Artificial Intelligence Review*, vol. 55, no. 7, pp. 5731–5780, Oct. 2022, doi: 10.1007/s10462-022-10144-1.

[12] O. Hamed and T. Zesch, "The role of diacritics in designing lexical recognition tests for Arabic," *Procedia Computer Science*, vol. 117, pp. 119–128, 2017, doi: 10.1016/j.procs.2017.10.100.

[13] H. Mubarak, A. Abdelali, K. Darwish, M. Eldesouki, Y. Samih, and H. Sajjad, "A system for diacritizing four varieties of Arabic," in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP): System Demonstrations*, 2019, pp. 217–222, doi: 10.18653/v1/D19-3037.

[14] M. N. Al-Kabi, N. A. Abdulla, and M. Al-Ayyoub, "An analytical study of Arabic sentiments: Maktoob case study," in *2013 IEEE Third International Conference on Information Science and Technology (ICIST)*, Dec. 2013, pp. 89–94, doi: 10.1109/ICIST.2013.6747511.

[15] S. Larabi Marie-Sainte, N. Alalyani, S. Alotaibi, S. Ghouzali, and I. Abunadi, "Arabic natural language processing and machine learning-based systems," *IEEE Access*, vol. 7, pp. 7011–7020, 2019, doi: 10.1109/ACCESS.2018.2890076.

[16] H. Saleh, S. Mostafa, L. A. Gabralla, A. O. Aseeri, and S. El-Sappagh, "Enhanced Arabic sentiment analysis using a novel stacking ensemble of hybrid and deep learning models," *Applied Sciences*, vol. 12, no. 18, p. 8967, Sep. 2022, doi: 10.3390/app12188967.

[17] H. K. Omar, M. Frikha, and A. K. Jumaa, "Improving big data recommendation system performance using NLP techniques with multi attributes," *Informatica*, vol. 48, no. 5, pp. 63–70, Feb. 2024, doi: 10.31449/inf.v48i5.5255.

[18] M. Al-Ayyoub, A. A. Khamaiseh, Y. Jararweh, and M. N. Al-Kabi, "A comprehensive survey of arabic sentiment analysis," *Information Processing & Management*, vol. 56, no. 2, pp. 320–342, Mar. 2019, doi: 10.1016/j.ipm.2018.07.006.

[19] A. M. Alayba, V. Palade, M. England, and R. Iqbal, "Improving sentiment analysis in Arabic using word representation," in *2018 IEEE 2nd International Workshop on Arabic and Derived Script Analysis and Recognition (ASAR)*, Mar. 2018, pp. 13–18, doi: 10.1109/ASAR.2018.8480191.

[20] A. Elnagar, L. Lulu, and O. Einea, "An Annotated Huge Dataset for Standard and Colloquial Arabic Reviews for Subjective Sentiment Analysis," *Procedia Computer Science*, vol. 142, pp. 182–189, 2018, doi: 10.1016/j.procs.2018.10.474.

[21] R. M. K. Saeed, S. Rady, and T. F. Gharib, "An ensemble approach for spam detection in Arabic opinion texts," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 1, pp. 1407–1416, Jan. 2022, doi: 10.1016/j.jksuci.2019.10.002.

[22] A. S. Salıhı and O. Yıldız, "Comparative Analysis of Ad Click Behavior Prediction Using GAN-Augmented Data and Traditional Machine Learning Techniques," *Politeknik Dergisi*, vol. 28, no. 2, pp. 683–695, Mar. 2025, doi: 10.2339/politeknik.1525138.

[23] X. Deng, Y. Li, J. Weng, and J. Zhang, "Feature selection for text classification: A review," *Multimedia Tools and Applications*, vol. 78, no. 3, pp. 3797–3816, Feb. 2019, doi: 10.1007/s11042-018-6083-5.

[24] Q. Liu, J. Wang, D. Zhang, Y. Yang, and N. Wang, "Text Features Extraction based on TF-IDF Associating Semantic," in *2018 IEEE 4th International Conference on Computer and Communications (ICCC)*, Dec. 2018, pp. 2338–2343, doi: 10.1109/CompComm.2018.8780663.

[25] T. Yu, C. Hidey, O. Rambow, and K. McKeown, "Leveraging Sparse and Dense Feature Combinations for Sentiment Classification," *Arxiv*, 2017, [Online]. Available: http://arxiv.org/abs/1708.03940.

[26] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *1st International Conference on Learning Representations, ICLR 2013 - Workshop Track Proceedings*, 2013.

[27] S. Kaddoura, G. Chandrasekaran, D. Elena Popescu, and J. H. Duraisamy, "A systematic literature review on spam content detection and classification," *PeerJ Computer Science*, vol. 8, p. e830, Jan. 2022, doi: 10.7717/peerj-cs.830.

[28] S. Areed, O. Alqaryouti, B. Siyam, and K. Shaalan, "Aspect-Based Sentiment Analysis for Arabic Government Reviews," in *Studies in Computational Intelligence*, vol. 874, 2020, pp. 143–162.

[29] A. A. Al Shamsi and S. Abdallah, "Ensemble Stacking Model for Sentiment Analysis of Emirati and Arabic Dialects," *Journal of King Saud University - Computer and Information Sciences*, vol. 35, no. 8, p. 101691, Sep. 2023, doi: 10.1016/j.jksuci.2023.101691.

[30] A. Alsanad, "An Improved Arabic Sentiment Analysis Approach using Optimized Multinomial Naïve Bayes Classifier," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 8, pp. 90–98, 2022, doi: 10.14569/IJACSA.2022.0130812.

[31] L. Moudjari, K. Akli-Astouati, and F. Benamara, "An algerian corpus and an annotation platform for opinion and emotion analysis," in *LREC 2020 - 12th International Conference on Language Resources and Evaluation, Conference Proceedings*, 2020, pp. 1202–1210.

[32] D. Ghoul, J. Patrix, G. Lejeune, and J. Verny, "A combined AraBERT and Voting Ensemble classifier model for Arabic sentiment analysis," *Natural Language Processing Journal*, vol. 8, p. 100100, Sep. 2024, doi: 10.1016/j.nlp.2024.100100.

[33] M. A. Galal, A. Hassan Yousef, H. H. Zayed, and W. Medhat, "Arabic sarcasm detection: An enhanced fine-tuned language model approach," *Ain Shams Engineering Journal*, vol. 15, no. 6, p. 102736, Jun. 2024, doi: 10.1016/j.asej.2024.102736.

[34] Q. Abuein, R. M. Al-Khatib, A. Migdady, M. S. Jawarneh, and A. Al-Khateeb, "ArSa-Tweets: a novel Arabic sarcasm detection system based on deep learning model," *Heliyon*, vol. 10, no. 17, p. e36892, Sep. 2024, doi: 10.1016/j.heliyon.2024.e36892.

[35] R. Bensoltane and T. Zaki, "Neural multi-task learning for end-to-end Arabic aspect-based sentiment analysis," *Computer Speech & Language*, vol. 89, p. 101683, Jan. 2025, doi: 10.1016/j.csl.2024.101683.

[36] Y. Matrane, F. Benabbou, and Z. Banou, "WeVoTe: a weighted voting technique for automatic sentiment annotation of Moroccan dialect comments," *IEEE Access*, vol. 12, pp. 16276–16298, 2024, doi: 10.1109/ACCESS.2024.3359430.

[37] A. Helmy, R. Nassar, and N. Ramdan, "Depression detection for twitter users using sentiment analysis in English and Arabic tweets," *Artificial Intelligence in Medicine*, vol. 147, p. 102716, Jan. 2024, doi: 10.1016/j.artmed.2023.102716.

[38] O. Alsemaree, A. S. Alam, S. S. Gill, and S. Uhlig, "Sentiment analysis of Arabic social media texts: a machine learning approach to deciphering customer perceptions," *Heliyon*, vol. 10, no. 9, p. e27863, May 2024, doi: 10.1016/j.heliyon.2024.e27863.

[39] R. Alrefae, R. Alqahmi, M. Alduraibi, S. Almatrafi, and A. Alayed, "Enhancing customer experience through arabic aspect-based sentiment analysis of Saudi reviews," *International Journal of Advanced Computer Science and Applications*, vol. 15, no. 7, pp. 421–427, 2024, doi: 10.14569/IJACSA.2024.0150742.

[40] M. Ramzy and B. Ibrahim, "User satisfaction with Arabic COVID-19 apps: sentiment analysis of users' reviews using machine learning techniques," *Information Processing & Management*, vol. 61, no. 3, p. 103644, May 2024, doi: 10.1016/j.ipm.2024.103644.

[41] F. Husain, H. Alostad, and H. Omar, "Bridging the Kuwaiti dialect gap in natural language processing," *IEEE Access*, vol. 12, pp. 27709–27722, 2024, doi: 10.1109/ACCESS.2024.3364367.

[42] T. M. Omran, B. T. Sharef, C. Grosan, and Y. Li, "Transfer learning and sentiment analysis of Bahraini dialects sequential text data using multilingual deep learning approach," *SSRN Electronic Journal*, 2022, doi: 10.2139/ssrn.4111929.

[43] S. F. Sabbeh and H. A. Fasihuddin, "A comparative analysis of word embedding and deep learning for Arabic sentiment classification," *Electronics*, vol. 12, no. 6, p. 1425, Mar. 2023, doi: 10.3390/electronics12061425.

[44] N. Elhassan *et al.*, "Arabic sentiment analysis based on word embeddings and deep learning," *Computers*, vol. 12, no. 6, p. 126, Jun. 2023, doi: 10.3390/computers12060126.

[45] A. Alqarni and A. Rahman, "Arabic tweets-based sentiment analysis to investigate the impact of COVID-19 in KSA: a deep learning approach," *Big Data and Cognitive Computing*, vol. 7, no. 1, p. 16, Jan. 2023, doi: 10.3390/bdcc7010016.

[46] A. M. Mostafa, M. Aljasir, M. Alruily, A. Alsayat, and M. Ezz, "Innovative forward fusion feature selection algorithm for sentiment analysis using supervised classification," *Applied Sciences*, vol. 13, no. 4, p. 2074, Feb. 2023, doi: 10.3390/app13042074.

[47] S. M. A. H. Shah, S. F. H. Shah, A. Ullah, A. Rizwan, G. Atteia, and M. Alabdulhafith, "Arabic sentiment analysis and sarcasm detection using probabilistic projections-based variational switch transformer," *IEEE Access*, vol. 11, pp. 67865–67881, 2023, doi: 10.1109/ACCESS.2023.3289715.

[48] M. E. M. Abo *et al.*, "A multi-criteria approach for Arabic dialect sentiment analysis for online reviews: exploiting optimal machine learning algorithm selection," *Sustainability*, vol. 13, no. 18, p. 10018, Sep. 2021, doi: 10.3390/su131810018.

[49] M. Hadwan, M. Al-Sarem, F. Saeed, and M. A. Al-Hagery, "An improved sentiment classification approach for measuring user satisfaction toward governmental services' mobile apps using machine learning methods with feature engineering and SMOTE technique," *Applied Sciences*, vol. 12, no. 11, p. 5547, May 2022, doi: 10.3390/app12115547.

[50] M. Hadwan, M. A. Al-Hagery, M. Al-Sarem, and F. Saeed, "Arabic sentiment analysis of users' opinions of governmental mobile applications," *Computers, Materials & Continua*, vol. 72, no. 3, pp. 4675–4689, 2022, doi: 10.32604/cmc.2022.027311.

[51] S. Al-Hagree and G. Al-Gaphari, "Arabic sentiment analysis based machine learning for measuring user satisfaction with banking services' mobile applications : comparative study," in *2022 2nd International Conference on Emerging Smart Technologies and Applications (eSmarTA)*, Oct. 2022, pp. 1–4, doi: 10.1109/eSmarTA56775.2022.9935486.

[52] R. Baly, A. Khaddaj, H. Hajj, W. El-Hajj, and K. B. Shaban, "ArSentD-LEV: a multi-topic corpus for target-based sentiment analysis in Arabic Levantine Tweets," *arXiv preprint arXiv:1906.01830*, 2019, [Online]. Available: http://arxiv.org/abs/1906.01830.

[53] "Feature selection optimization for highlighting opinions using supervised and unsupervised learning on arabic language," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 10, no. 2, pp. 636–642, Apr. 2021, doi: 10.30534/ijatcse/2021/251022021.

[54] M. Kasri, M. Birjali, and A. Beni-Hssane, "A comparison of features extraction methods for Arabic sentiment analysis," in *Proceedings of the 4th International Conference on Big Data and Internet of Things*, Oct. 2019, pp. 1–6, doi: 10.1145/3372938.3372998.

[55] B. Talafha *et al.*, "Multi-dialect arabic BERT for country-level dialect identification," *arXiv preprint arXiv:2007.05612*, 2020, [Online]. Available: http://arxiv.org/abs/2007.05612.

[56] J. O. Atoum and M. Nouman, "Sentiment analysis of Arabic Jordanian dialect tweets," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 2, pp. 256–262, 2019, doi: 10.14569/IJACSA.2019.0100234.

[57] S. Alosaimi, M. Alharthi, K. Alghamdi, T. Alsubait, and T. Alqurashi, "Sentiment analysis of arabic reviews for saudi hotels using unsupervised machine learning," *Journal of Computer Science*, vol. 16, no. 9, pp. 1258–1267, Sep. 2020, doi: 10.3844/jcssp.2020.1258.1267.

[58] H. E. Elagbry, S. Attia, A. Abdel-Rahman, A. Abdel-Ate, and S. Girgis, "A contextual word embedding for Arabic sarcasm detection with random forests," in *WANLP 2021 - 6th Arabic Natural Language Processing Workshop, Proceedings of the Workshop*, 2021, pp. 340–344.

[59] A. El Mekki, A. El Mahdaouy, I. Berrada, and A. Khoumsi, "Domain adaptation for Arabic cross-domain and cross-dialect sentiment analysis from contextualized word embedding," in *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2021, pp. 2824–2837, doi: 10.18653/v1/2021.naacl-main.226.

[60] A. A. Al-Futamani and H. H. Al-Baity, "Emotional analysis of Arabic Saudi dialect tweets using a supervised learning approach," *Intelligent Automation & Soft Computing*, vol. 29, no. 1, pp. 89–109, 2021, doi: 10.32604/iasc.2021.016555.

[61] B. Brahimi, M. Touahria, and A. Tari, "Improving sentiment analysis in Arabic: a combined approach," *Journal of King Saud University - Computer and Information Sciences*, vol. 33, no. 10, pp. 1242–1250, Dec. 2021, doi: 10.1016/j.jksuci.2019.07.011.

[62] L. Moudjari, F. Benamara, and K. Akli-Astouati, "Multi-level embeddings for processing Arabic social media contents," *Computer Speech & Language*, vol. 70, p. 101240, Nov. 2021, doi: 10.1016/j.csl.2021.101240.

[63] M. Alruily and O. R. Shahin, "Sentiment analysis of Twitter data for Saudi Universities," *International Journal of Machine Learning and Computing*, vol. 10, no. 1, pp. 18–24, 2020, doi: 10.18178/ijmlc.2020.10.1.892.

[64] E. Boujou, H. Chataoui, A. El Mekki, S. Benjelloun, I. Chairi, and I. Berrada, "An open access NLP dataset for Arabic dialects : Data collection, labeling, and model construction," *arXiv preprint arXiv:2102.11000*, 2021, [Online]. Available: http://arxiv.org/abs/2102.11000.

[65] H. Elfaik and E. H. Nfaoui, "A comparative evaluation of classification algorithms for sentiment analysis using word embeddings," in *Advances in Intelligent Systems and Computing*, vol. 1105 AISC, 2020, pp. 1–11.

[66] I. Kaibi, E. H. Nfaoui, and H. Satori, "Sentiment analysis approach based on combination of word embedding techniques," in *Advances in Intelligent Systems and Computing*, vol. 1076, 2020, pp. 805–813.

[67] A. A. Hnaif, E. Kanan, and T. Kanan, "Sentiment analysis for Arabic social media news polarity," *Intelligent Automation & Soft Computing*, vol. 28, no. 1, pp. 107–119, 2021, doi: 10.32604/iasc.2021.015939.

[68] A. El kah and I. Zeroual, "Sentiment analysis of Facebook users towards COVID-19 vaccination," in *Proceedings of the 6th International Conference on Networking, Intelligent Systems & Security*, May 2023, pp. 1–4, doi: 10.1145/3607720.3607753.

[69] D. Wehbe, A. Alhammadi, H. Almaskari, K. Alsereidi, and H. Ismail, "UAE e-Learning sentiment analysis framework," in *The 7th Annual International Conference on Arab Women in Computing in Conjunction with the 2nd Forum of Women in Research*, Aug. 2021, pp. 1–4, doi: 10.1145/3485557.3485570.

[70] S. El Ouahabi, S. El Ouahabi, and E. W. Dadi, "Multi-domain dataset for moroccan arabic dialect sentiment analysis in social networks," in *2023 IEEE 6th International Conference on Cloud Computing and Artificial Intelligence: Technologies and Applications (CloudTech)*, Nov. 2023, pp. 01–07, doi: 10.1109/CloudTech58737.2023.10366119.

[71] B. Klouche, S. M. Benslimane, and N. Mahammed, "Sentiment analysis of algerian dialect using a deep learning approach," in *Lecture Notes in Networks and Systems*, vol. 413 LNNS, 2022, pp. 122–131.

[72] Y. Zahidi, Y. Al-Amrani, and Y. El Younoussi, "Improving Arabic sentiment analysis using LSTM based on word embedding models," *Vietnam Journal of Computer Science*, vol. 10, no. 03, pp. 391–407, Aug. 2023, doi: 10.1142/S2196888823500069.

[73] C. Fourati, H. Haddad, A. Messaoudi, M. B. H. Hmida, A. B. E. Mabrouk, and M. Naski, "Introducing a large Tunisian Arabizi dialectal dataset for sentiment analysis," in *WANLP 2021 - 6th Arabic Natural Language Processing Workshop, Proceedings of the Workshop*, 2021, pp. 226–230.

[74] S. M. Mohamed, E. H. Mohamed, and M. A. Belal, "Polarity detection of dialectal arabic using deep learning models," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 11, 2021, doi: 10.14569/IJACSA.2021.0121125.

[75] A. El Mekki, A. Alami, H. Alami, A. Khoumsi, and I. Berrada, "Weighted combination of BERT and N-GRAM features for Nuanced Arabic Dialect Identification," *Proceedings of the Fifth Arabic Natural Language Processing Workshop*, pp. 268–274, 2020.

[76] A. M. Bdeir and F. Ibrahim, "A framework for Arabic Tweets multi-label classification using word embedding and neural networks algorithms," in *Proceedings of the 2020 2nd International Conference on Big Data Engineering*, May 2020, pp. 105–112, doi: 10.1145/3404512.3404526.

[77] N. Azzouza, K. Akli-Astouati, and R. Ibrahim, "TwitterBERT: framework for Twitter sentiment analysis based on pre-trained language model representations," in *Advances in Intelligent Systems and Computing*, vol. 1073, 2020, pp. 428–437.

[78] H. Chouikhi, H. Chniter, and F. Jarray, "Arabic sentiment analysis using BERT model," in *Communications in Computer and Information Science*, vol. 1463, 2021, pp. 621–632.

[79] A. Abdelali, S. Hassan, H. Mubarak, K. Darwish, and Y. Samih, "Pre-Training BERT on Arabic Tweets: practical considerations," *arXiv preprint arXiv:2102.10684*, 2021, [Online]. Available: http://arxiv.org/abs/2102.10684.

[80] A. Masmoudi, J. Hamdi, and L. H. Belguith, "Deep learning for sentiment analysis of tunisian dialect," *Computacion y Sistemas*, vol. 25, no. 1, pp. 129–148, 2021, doi: 10.13053/CYS-25-1-3472.

[81] Z. Benmounah, A. Boulesnane, A. Fadheli, and M. Khial, "Sentiment analysis on algerian dialect with transformers," *Applied Sciences*, vol. 13, no. 20, p. 11157, Oct. 2023, doi: 10.3390/app132011157.

[82] M. Mhamed *et al.*, "A deep CNN architecture with novel pooling layer applied to two Sudanese Arabic sentiment data sets," *Journal of Information Science*, Oct. 2023, doi: 10.1177/01655515231188341.

[83] 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," *Computational Intelligence and Neuroscience*, vol. 2021, no. 1, Jan. 2021, doi: 10.1155/2021/5538791.

[84] M. Berrimi, M. Oussalah, A. Moussaoui, and M. Saidi, "Attention mechanism architecture for arabic sentiment analysis," *ACM Transactions on Asian and Low-Resource Language Information Processing*, vol. 22, no. 4, pp. 1–26, Apr. 2023, doi: 10.1145/3578265.

[85] H. Chahidi, H. Omara, M. Lazaar, and M. Al Achhab, "Impact of neural network architectures on Arabic sentiment analysis," in *Proceedings of the 4th International Conference on Big Data and Internet of Things*, Oct. 2019, pp. 1–6, doi: 10.1145/3372938.3372950.

[86] M. Errami, M. A. Ouassil, R. Rachidi, B. Cherradi, S. Hamida, and A. Raihani, "Investigating the performance of BERT model for sentiment analysis on moroccan news comments," in *2023 3rd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)*, May 2023, pp. 1–8, doi: 10.1109/IRASET57153.2023.10152965.

[87] I. Guellil *et al.*, "A semi-supervised approach for sentiment analysis of Arab(ic+izi) Messages: Application to the Algerian Dialect," *SN Computer Science*, vol. 2, no. 2, p. 118, Apr. 2021, doi: 10.1007/s42979-021-00510-1.

[88] A. Oussous, F.-Z. Benjelloun, A. A. Lahcen, and S. Belfkih, "ASA: a framework for Arabic sentiment analysis," *Journal of Information Science*, vol. 46, no. 4, pp. 544–559, Aug. 2020, doi: 10.1177/0165551519849516.

[89] Y. Matrane, F. Benabbou, and N. Sael, "Sentiment analysis through word embedding using AraBERT: Moroccan dialect use case," in *2021 International Conference on Digital Age & Technological Advances for Sustainable Development (ICDATA)*, Jun. 2021, pp. 80–87, doi: 10.1109/ICDATA52997.2021.00024.

[90] I. Abu Farha and W. Magdy, "A comparative study of effective approaches for Arabic sentiment analysis," *Information Processing & Management*, vol. 58, no. 2, p. 102438, Mar. 2021, doi: 10.1016/j.ipm.2020.102438.

[91] O. Elsamadony, A. Keshk, and A. Abdelatey, "Sentiment analysis for Arabic language using word embedding," in *2021 17th International Computer Engineering Conference (ICENCO)*, Dec. 2021, pp. 51–56, doi: 10.1109/ICENCO49852.2021.9698960.

[92] A. C. Mazari and A. Djeffal, "Sentiment analysis of Algerian dialect using machine learning and deep learning with Word2vec," *Informatica*, vol. 46, no. 6, pp. 67–78, Jul. 2022, doi: 10.31449/inf.v46i6.3340.

[93] I. Guellil, M. Mendoza, and F. Azouaou, "Arabic dialect sentiment analysis with ZERO effort. \\ Case study: Algerian dialect," *Inteligencia Artificial*, vol. 23, no. 65, pp. 124–135, Jul. 2020, doi: 10.4114/intartif.vol23iss65pp124-135.

[94] K. Hamadouche, K. Z. Bousmaha, M. A. Bekkoucha, and L. Hadrich-Belguith, "AlgBERT: automatic construction of annotated corpus for sentiment analysis in Algerian dialect," *ACM Transactions on Asian and Low-Resource Language Information Processing*, vol. 22, no. 12, pp. 1–17, Dec. 2023, doi: 10.1145/3632948.

[95] D. S. Abdelminaam, N. Neggaz, I. A. E. Gomaa, F. H. Ismail, and A. A. Elsawy, "ArabicDialects: an efficient framework for arabic dialects opinion mining on Twitter using optimized deep neural networks," *IEEE Access*, vol. 9, pp. 97079–97099, 2021, doi: 10.1109/ACCESS.2021.3094173.

[96] N. Habbat, H. Nouri, H. Anoun, and L. Hassouni, "Sentiment analysis of imbalanced datasets using BERT and ensemble stacking for deep learning," *Engineering Applications of Artificial Intelligence*, vol. 126, p. 106999, Nov. 2023, doi: 10.1016/j.engappai.2023.106999.

[97] Y. Khalifa and A. Elnagar, "Sentiment analysis of colloquial Arabic Tweets with Emojis," in *Advances in Intelligent Systems and Computing*, vol. 1339, 2021, pp. 418–430.


## APPENDIX

Table 1. Overview of research papers included in the review

| Resource | Years | Data source | Size of datasets | Dialect | Feature extractor | Classifier | Best performance | Accuracy |
|---|---|---|---|---|---|---|---|---|
| [28] | 2020 | Twitter | 9,000 | Algerian dialect | BOW | BERT, CNN | BERT | 82% |
| [29] | 2024 | Twitter | 55000 | Multi | TF-IDF | AraBERT | AraBERT + Voting Ensemble | 81.34% |
| [30] | 2024 | Twitter | 26014 | Multi | N/A | DNN, AraBert | AraBert | 84.29% |
| [31] | 2024 | Twitter | 3000 | Multi | TF-IDF, word embedding | LSTM, Multi-headed CNN-LSTM-GRU, BERT, AraBert-V01, AraBert-V02 | AraBert-V02 | 87.00% |
| [32] | 2024 | Twitter | 10000 | Multi | TF-IDF, BOW | Lgbm, RF, SVM, Logistic Regression (LR) | svm | F1-scor = 96.6% |
| [33] | 2024 | Hotel reviews | 15,562 | Multi | Bert | BERT, BiGRU, CRF | BERT-BiGRU-ATT-CRF | F1-scor = 80.36 % |

Table 1. Overview of research papers included in the review *(Continued)*

| Resource | Years | Data source | Size of datasets | Dialect | Feature extractor | Classifier | Best performance | Accuracy |
|---|---|---|---|---|---|---|---|---|
| [34] | 2024 | Twitter | 10,646 | Saudi dialect | TF-IDF, minimum redundancy maximum relevance (MRMR) | KNN, SVM, DT, RF | SVM, DT | 95.00% |
| [35] | 2024 | Google Play Store | 114,499 | Multi | N-grams, POS | SVM, KNN, NB, LR, RF, NN | ANN | 89.00% |
| [36] | 2024 | Mix | 24,978 | Moroccan dialect | AraBert | BiGRU, CNN-BiGRU, LSTM, CNN-LSTM, BiLSTM, CNN-BiLSTM | (AraBert+BiGRU, AraBert+CNN-BiGRU) | 91.25% |
| [37] | 2024 | Mix | 4,000 | Saudi dialect | POS, TF-IDF, n-grams, | SVM, RF, NB, KNN, | SVM | 86.51% |
| [38] | 2023 | Amazon reviews | 5000 | Bahraini dialect | BERT | LSTM, | LSTM | 96.72% |
| [39] | 2023 | Twitter | 55,000 | Multi | TF–IDF | LR, NB, SVM, RF, XGB, DT, KNN | NB | 84.40% |
| [40] | 2023 | Mix | 237,800 | Multi | GloVe, Word2vec, FastText and ARBERT | BiLSTM and CNN | BERT + BiLSTM | 93.97% |
| [41] | 2024 | Twitter | 16,600 | Kuwaiti dialect | TF-IDF | LR, SVM, MNB, AraBert | AraBert | 87.00% |
| [42] | 2023 | Hotel reviews | 436,772 | Multi | Word2Vec and fastText | CNN, LSTM, CNN-LSTM | CNN | 94.69% |
| [43] | 2023 | Twitter | 157,214 | Saudi dialect | Word2Vec and fastText | BiLSTM, CNN | CNN | 92.80% |
| [44] | 2023 | Twitter | 27,000 | Multi | N/A | RF, KNN, NB, DT, LR, SVM, proposed (MST) | proposed (MST) | 52.00% |
| [45] | 2021 | Twitter | 11647 | Saudi dialect | N/A | SVM, KNN, NB | SVM | 85.25 |
| [46] | 2022 | Twitter | 51000 | Saudi dialect | BOW, TF-IDF, Word2Vec | SVM, RF, LR, NB | NB | 74.25% |
| [47] | 2022 | Mix | 8000 | Saudi dialect | TDM | DT, SVM, NB, KNN | KNN | 78.46%. |
| [48] | 2022 | Google Play Store | 3545 | Yemen dialect | TF-IDF | NB, KNN, DT, SVM, | NB | 89.65% |
| [49] | 2020 | Twitter | 4,00 | Multi | n-grams, TF-IDF | SVM, RF, LR | LR | 63% |
| [50] | 2021 | Mix | 3,015 | Modern Arabic standard | TFMX, TDMX | NBC, KNN, k-means, fuzzy c-means | KNN | 83.60% |
| [51] | 2020 | Facebook | 10,254 | Moroccan dialect | BOW, TF-IDF, word2vec | LR, SVM, RF, Extra Trees | Logistic Regression+ (TD-IDF) | 82.10% |
| [52] | 2020 | Twitter | 31,000 | Multi | N/A | BERT | BERT | 45.07% |
| [53] | 2020 | Twitter | 1000 | Jordanian dialect | N-gram | SVM, NB | SVM | 77% |
| [54] | 2020 | TripAdvisor | 4604 | Saudi dialect | TF-IDF | k-means, Support Vector Clustering | k-means clustering | 76% |

Table 1. Overview of research papers included in the review *(Continued)*

| Resource | Years | Data source | Size of datasets | Dialect | Feature extractor | Classifier | Best performance | Accuracy |
|---|---|---|---|---|---|---|---|---|
| [55] | 2021 | Facebook | 255,008 | MSA, Algerian dialect | Word2vec, FastText | RF, LR, CNN, LSTM | CNN | 89% |
| [56] | 2020 | Facebook, Twitter, YouTube | 2000 | Moroccan dialect | N/A | NB, SVM, LSTM, CNN | CNN | 96% |
| [57] | 2021 | Twitter | 12548 | Multi | N/A | RF | RF | 75% |
| [58] | 2021 | Mix | 353, 171 | Multi | Skip-gram, BOW using | Bert | Bert | 91.5 |
| [59] | 2021 | Twitter | 5,149 | Saudi dialect | N-gram, chi-square statistical | LR, SVM, MNB | LR, SVM | 73.39% |
| [60] | 2023 | Twitter | 52155 | Multi | Word2Vec, FastText | LSTM | LSTM+FastText | 84,04 |
| [61] | 2021 | Arabic reviews | 500 | MSA | n-gram, skip-n-gram, POS | NB, Maximum Entropy (MA), SVM | SVM + n-gram | 94.20% |
| [62] | 2021 | Twitter | 2,236,351 | Algerian dialect | Word2Vec, CBOW, FastText | CNN | EmbCNN | 83.63 |
| [63] | 2021 | Twitter | 100000 | Tunisian dialect | BERT | MBERT | MBERT | 82.90% |
| [64] | 2021 | Twitter | 40000 | Egyptian dialect | CBOW | CNN, LSTM, BiLSTM, | CNN | 99.65% |
| [65] | 2020 | Twitter | 60000 | Saudi dialect | Word2Vec, TF-IDF, TDM | SVM, NB, KNN, RF | SVM | 93.529 |
| [66] | 2021 | Twitter | 12000 | Moroccan dialect | AraBert | SVM, NB, DT, LR, NN, LSTM, BLSTM | BLSTM | 83.24% |
| [67] | 2020 | social media | 30957 | Multi | N-gram, TF–IDF | Bert | Bert+(TF-IDF) | 40.95% |
| [68] | 2021 | social media | 50000 | Multi | BOW, TFIDF | SGD Classifier, NB, LSVM | SGD Classifier | 77.00% |
| [69] | 2020 | Twitter | 160000 | MSA | Word2Vec | CNN, RNN | CNN | 90.15% |
| [70] | 2020 | Twitter | 1387085 | Multi | Word2Vec | CNN, BLSTM, BERT | BERT | 71.96% |
| [37] | 2020 | Mix | 61582 | Multi | CBOW | NB, SVM, LR, SGD, RF, KNN, DT, AdaBoost | LR | 81.68% |
| [71] | 2020 | Mix | 172500 | Multi | BERT | BERT | BERT | 96,11% |
| [72] | 2020 | Mix | 20314 | Multi | FastText, AraVec, | SVM, NuSVM, RF, LR, NB, SGD | NuSVM + Concatenation (AraVec, FastText) | 92.22% |
| [73] | 2021 | Mix | N/A | Multi | N/A | BERT | QARiB+ entity recognition | 69.10% |
| [74] | 2021 | Mix | 249941 286 | Egyptian dialect, MSA | TF-IDF, n-gram | Naive Bayes, SVM, BiLSTM, LSTM, CNN, AraBERT | AraBERT | 93.00% |
| [75] | 2021 | Arabic Book Reviews | 63000 | MSA | BOW, Skip-Gram, TD-IDF, | LR, KNN, SVM, RF, DT, XGBoost, CNN, RNN | (CNN-RNN) + Skip-gram | 88.38% |
| [76] | 2021 | FaceBook | 21000 | Tunisian dialect | Word2Vec | CNN, LSTM, BiLSTM | LSTM, BiLSTM | 87.00% |
| [77] | 2023 | YouTube comments | 45000 | Algerian Dialect | n-grams | BERT, LSTM, Bi-LSTM | BERT | 81.74% |
| [78] | 2022 | Mix | 11760 | Algerian dialect | BOW, Skip-Gram) | SVM, NB, MNB, CNN, RNN | MNB | 84.21% |

Table 1. Overview of research papers included in the review *(Continued)*

| Resource | Years | Data source | Size of datasets | Dialect | Feature extractor | Classifier | Best performance | Accuracy |
|---|---|---|---|---|---|---|---|---|
| [79] | 2022 | Twitter, FaceBook | 11109 | Sudanese dialect | TD-IDF | LR, RF, NB, SVM, RNN, CNN, CNN+LSTM, SCM+MMA | SCM+MMA | 88.37% |
| [80] | 2021 | Twitter | 10106919 | Saudi dialect | Word2Vec, AraVec | MC1 (CNN+Global Average Pooling), MC2 (CNN + bidirectional gated recurrent units) | MC2 | 74.35% |
| [81] | 2020 | FaceBook | 500 | Algerian dialect | word2Vec, BOW, Doc2Vec, FastText | NB, LR, RF, SGD, SVC, CNN, MLP, LSTM, BiLSTM | SGD(Word2Vec) | 89.00% |
| [82] | 2023 | Twitter, YouTube | 54000 | Algerian, Standard Arabic | CBOW, FastText, | RF, SVM, GNB, MNB, AlgBERT, CNN-LSTM | AlgBERT | 92.62% |
| [83] | 2021 | Facebook, Twitter | 6138 | MSA | BOW | SVM, NB, KNN, RF, DT | SVM | 83.00% |
| [84] | 2023 | Book, Hotel reviews | 679555 | MSA | FastText, Learnable embedding | Bi-GRU, Bi-LSTM | GRU/LSTM + Leranable embadding | 95.65% |
| [85] | 2020 | Twitter | 2000 | MSA | TF-IDF, | Multi-Layer Perceptron | Multi-Layer Perceptron | 85.28% |
| [86] | 2023 | FaceBook | 2600 | Moroccan dialect | TF-IDF | NB, SVM, RF | RF | 87.50% |
| [87] | 2021 | Twitter | 171873 | UAE dialect | BOW | DT, RF, GB | RF | 88.72% |
| [88] | 2023 | HESPRESS | 14050 | Moroccan dialect | Arabic Bert | MARABERT, AraBERT | MARABERT | 92.31% |
| [89] | 2023 | Mix | 70.332 | Moroccan dialect | BOW | DT, MNB, SVM, RF, LR | LR | 75.00% |
| [90] | 2021 | Mix | 5684700 | Multi | TF-IDF, Skip-gram | NB, KNN, LT, SVM, DT, RF, ELSTM, GRU, mLSTM | mLSTM | 97.67% |
| [91] | 2024 | Twitter | 57391 | MSA, Egyptian, Gulf | TF-IDF, BOW | LGBM, RF, SVM, LR | SVM+BOW | 96.40% |
| [92] | 2022 | Facebook | 65,125 | MSA, Algerian dialect | TF-IDF | SVM, CNN | CNN | 74.66% |
| [93] | 2023 | Mix | 171,700 | Moroccan dialect, MSA | BERT | SVM, LSTM, BiLSTM, GRU | stacking ensemble (using SVM) | 93.20% |
| [94] | 2021 | Mix | 1,032,284 | Egyptian dialect and MSA | TF-IDF, FastText, GloVe | GB, LR, AdaBoost, DT, MNB, SVM, XGB, RF, MLP, CNN | CNN+ FastText | 89.00% |
| [95] | 2023 | Twitter | 50000 | Moroccan dialect | AraBERT, Qarib | SVM, AraBERT, Qarib | Qarib+ Qarib | 93.00% |
| [96] | 2023 | Facebook | 9,396 | Tunisian dialect | GloVe, TF-IDF, BOW | NB, SVM, RF, LR, DT, CNN, LSTM, Bi-LSTM, RNN | Bi-LSTM | 79,80% |
| [97] | 2022 | Twitter | 199,000 | Saudi dialect | TF-IDF, BOW | SVM, NB, LR, DT, RF, XGBoost, RNN, LSTM, Bi-LSTM, CNN | Bi-LSTM + Word Embedding | 87.00% |

**BIOGRAPHIES OF AUTHORS**

**Abbas Hussein Ali** received his B.Sc. degree in Computer Science from Kirkuk University. He is currently working as an assistant lecturer at the Faculty of Computer Science and Information Technology, Kirkuk University. He is also pursuing a Ph.D. degree in the Department of Computer Science, Graduate School of Informatics, Gazi University, Ankara, Turkey. His current research interests focus on Arabic natural language processing. He can be contacted at email: abbasalsaka@uokirkuk.edu.iq.

**Prof. Dr. Necaattin Barışçı** received the B.Sc. degree in Electronics and Computer Education from Gazi University in 1994. He completed his M.Sc. in Electronics and Computer Education at the Institute of Science, Gazi University, in 2000 and earned his Ph.D. from the same institution in 2005. He is currently the Head of the Computer Science Department at the Institute of Informatics, Gazi University. His research interests include AI, machine learning, pattern recognition, fuzzy systems, image processing, neural networks, and programming languages. He can be contacted at email: nbarisci@gazi.edu.tr.