Journal of Artificial Intelligence and Data Science (JAIDA)
e-ISSN: 2791-8335
https://dergipark.org.tr/pub/jaida

Vol. 2, No. 1, pp. 24-30, 2022
Review Article

Arabic Sentiment Analysis: Reviews of the Effective Used Algorithms

İnas CUMAOĞLU [1*], Vedat TÜMEN [2], Yüksel ÇELIK [1]

[1] Karabük University, Engineering Faculty, Department of Computer Engineering, Turkey
[2] Bitlis Eren University, Engineering Faculty, Department of Computer Engineering, Turkey

Abstract
Sentiment analysis (SA) attracted many researchers due to its success in many areas such as marketing, health, and politics. It is a science of artificial intelligence (AI) and natural language processing (NLP), which aims to study people's thoughts, attitudes, and aspirations on a subject. SA is based on textual data obtained from internet sites such as electronic stores, flight and hotel reservation sites, and social media sites like Twitter and Facebook. However, the problem with that data is that it is unstructured and unorganized. The researchers had to work on organizing it using NLP tools to deal with it and analyze the feelings extracted from this data. Due to the grammatical and morphological complexity of the Arabic language and the lack of an Arabic corpus, it is still in the early phases of processing Arabic texts compared to English texts. As a result, in this research, we examined the most recent literature and scientific studies on Arabic sentiment analysis (ASA) to identify the most important algorithms that have demonstrated their quality and usefulness in this sector. We observed the researchers' interest in the use of deep learning algorithms (DL), which demonstrated their efficacy in this field and the employment of a variety of text extraction approaches, the most prominent of which were the TF-IDF CBOW and Skip-gram.
Keywords: Arabic sentiment analysis; data mining, deep learning; NLP; social media.

1. Introduction

Sentiment analysis (SA) is one of the most important and most attractive sciences for researchers. Furthermore, it is one of the applications of natural language processing (NLP), which was born from three sciences, namely, human linguistics, computer science, and artificial intelligence [1]. Some researchers also call it data mining. The importance of this science lies in the ability to apply it to many areas of life, such as politics, health, marketing, finance, and others.

In marketing, it has been proven that SA should be applied to giant companies and institutions that wish to enhance their revenues and customer base by knowing the opinion of customers about their products and services. So the process of knowing the customer's opinion and reading his thoughts and aspirations is the process of SA. [2].

SA participated in political fields and predicted the victory of a candidate by analyzing the feelings of the people and the public [3]. In the financial markets, automated applications based on sentiment analysis have proven to be more efficient than human analysts[4]. Although sentiment analysis has made tangible progress in English texts, it is still in the process of being challenged and discovered in Arabic texts. The Arabic language is complex and rich in words, synonyms, and expressions, which made addressing it a challenge and a science racing to fill its gaps [5]. Moreover, because social media was overcrowded with Arab subscribers and users, the Arabic data became huge, large, and diverse. So researchers needed to collect this unstructured data by reorganizing it and using it in natural language processing and sentiment analysis in SA.

2. Sentiment Analysis

In recent years, social networking sites have increased and become more and more popular, which has led to the presence of extensive and big data in different languages, including the Arabic language, which has a diversity and multiplicity of dialects. Indeed, this data is helpful, but it is random and unorganized.

So, NLP and SA came to process these data, contribute to organizing it, and benefit from knowing opinions, feelings, and trends in various economic and political aspects. Such as improving product quality, predicting a candidate's election victory, forecasting stocks and the stock market, and various other applications.

SA is one of the NLP applications, as shown in Figure 1, concerned with studying people's attitudes, thoughts, and tendencies through comments, posts, reviews, or tweets that they share on websites and social media. Thus texts are divided into categories and classes using artificial intelligence AI, ML, and DL algorithms.

*Corresponding author e-mail address: inascu90@gmail.com

Received: 08/05/2022; Accepted: 22/06/2022

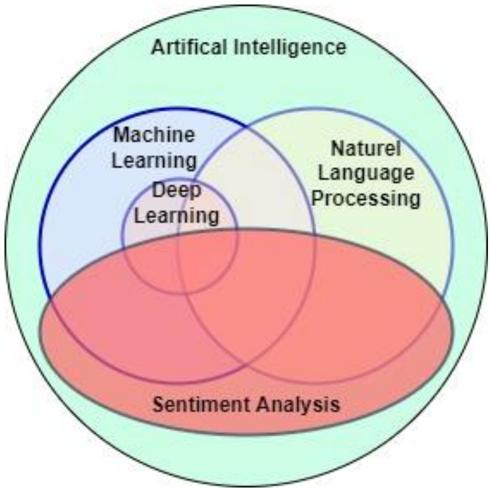

Figure 1. *Artificial Intelligence, Machine Learning, Deep Learning, Natural Language Processing, and Sentiment Analysis*

SA is also divided into three levels, as shown in Figure 2, where the document level deals with the feelings in the file as a whole without dividing it into sentences. In contrast, the sentence level deals with the feelings in each sentence separately, after making sure that it is a subject sentence containing feelings and not an object sentence without feelings and talking about facts. The aspect-based sentiment analysis level (ABSA) evaluates the feeling after dividing the content into several aspects and then discovers the feeling or polarity in each sentence after defining the aspect [6].

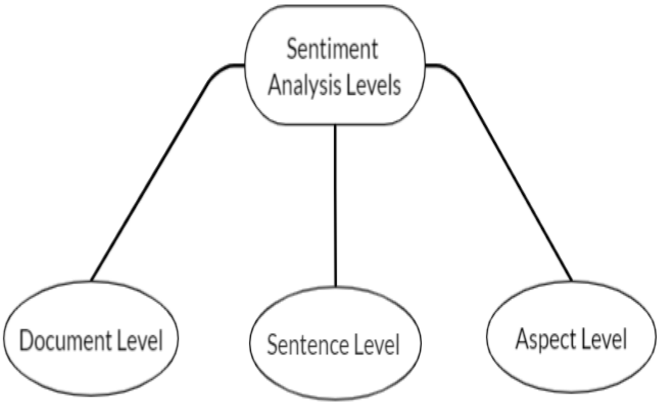

Figure 2. *Sentiment Analysis Levels*

2.1. Arabic sentiment analysis challenges

The Arabic language is a morphologically complex language with a plurality of dialects. The pioneers of social communication prefer to express their opinions using the various dialects of their countries. That has led to difficulty addressing the Arabic language because each dialect contains expressions that differ from country to country. For example, the word for a man in the Levant is "rajul"; In Sudan; it is "Zol " Also, the Arabic language does not contain uppercase or lowercase letters; instead, the Arabic letters differ in how they are written according to their position in the word. In addition, it is written from right to left, in contrast to writing the Latin letters. In addition, the diacritics of Arabic letters and prefixes and suffixes constitute an obstacle in the treatment of the Arabic language. Also, the lack of Arab corpus is one of the biggest challenges facing researchers [7]

3. Literature Review

In this section, we will introduce the recent literature with the used algorithms and the results of the experiments:

Alharbi and Qamar [8] suggested analyzing customers' feelings who visit cafes and restaurants in the Qassim region by conducting an opinion survey via Microsoft Form to improve the Saudi economy. A dataset with a size of 1785 reviews was created, and then it was reduced after filtering and pre-processing to 1,507 reviews in Arabic. Using five machine learning algorithms (KNN, NB, LR, AVM, and RF) after feature extraction by Term Frequency Inverse Document Frequency (TF-IDF), they could reach 89% accuracy by SVM, 93% F-measure by RF, and 92 % recall also by SVM.

In order to classify the text into two categories, positive and negative, Sayed et al. [9] created a dataset from the reviews of hotel customers on the Booking.com website. Thus, an Arabic corpus was created in the colloquial dialect and MSA with the name (RSAC), an abbreviation for Review Sentiment Analysis Corpus, and includes 6318 reviews divided into 3354 positive and 2964 negative. They used TF_IDF with unigrams after the initial processing, including stop words and non-Arabic letter removal. Then trained the model using nine algorithms of machine learning, six traditional (LR, SVM, KNN, NB, DT, RF) and three new ones not used previously in text classification (Gradient Boosting(GB), Ridge classifier (RC), Multilayer Perception (MCL)). At last, they tested the performance many times: one without applying any pre-processing and one more with applying stop words removal, another with the effect of stemming, and one with utilizing all pre-processing techniques according to accuracy, recall, precision, and F1- score and found that RC was the best one with 95% accuracy.

Abuuznien et al. [10] suggested using machine learning algorithms to know the Sudanese people's opinion of the ride-sharing service. So they collected Arabic corps from Twitter in their Sudanese dialect. The dataset consisted of 768 positive, 841 negatives, and 507 neutral tweets. It extracted 686 stop words in the Sudanese dialect added to the MSA stop word set. Then the pre-possessing phase started, and they used TF_IDF with n-gram to convert words into a vector, and in the classification phase, they used KNN, NB, and SVM to divide tweets into positive, negative, or neutral. They tested the efficiency of the models after following several techniques from Pre-Processing by measuring accuracy and F1 score, and they noticed that SVM presented a higher efficiency than the other two classifiers, with an accuracy of 95% after applying stemming as Pre-Processing.

Al-Tamimi et al. [11] collected data from comments on YouTube videos using YouTube API to analyze sentiments from the Arabic text. They also developed a windows application to help them annotate the data, so they built the dataset based on positive, negative, and neutral comments and the relevance of the comment to the content of the video. Then they started with the pre-processing step by removing comments written in a non-Arabic language and removing the diacritics, and using unigram and bigram with TF-IDF to create the input vector. Using KNN, Bernoulli NB, and SVM with RBF (Radial Bases Function), the data were classified in various forms (balanced, unbalanced, raw, after normalization, with positive and negative classes, and the dataset with three classes positive, negative and neutral). Consequently, they found that SVM with RBF achieved the best result in terms of F-Measure when they applied it to the unbalanced two-class dataset and after the normalization process.

In a case study of the opinions and sentiments of Qassim University students, Alassaf et al. [12] suggested the use of ABSA. They started collecting data related to Al Qassim University from Twitter via Twitter API, and it was 8,234 tweets in size. It was reduced to 7,934 after deleting the tweets whose classification differed by annotators and after the Pre-Processing phase. In this work, they chose nine aspects for analysis in aspect detection, including the educational Aspect, the activities, and the environment. In addition, they defined the polarity of opinion in the Aspect- opinion classification task, which they suggested calling the classes negative or non-negative polarity. The non-negative category included a few positive tweets with neutral tweets. They used SVM with ANOVA to determine features and determine the affiliation of the tweet to one of the nine aspects. They also used F- score 20% _ Reg technique to reduce features to be 7,361 after it was 161,396 (4.5% decreased). Then they utilized SVM and cross-validation with ten fold to measure the model's efficiency.

Omara et al. integrated 13 available Arabic datasets in Modern Standard Arabic (MSA) and Dialectal from various domains, including book, film, restaurant, hotel, product, and tweet reviews. There are 92123 items in the new dataset, with 71112 positive and 21011 positive samples. Pre-processing was used to increase the quality of the results. For Arabic sentiment analysis (ASA), they proposed employing three deep convolutional neural networks (CNNs). Moreover, to extract features, they used character level with CNN. Hence, they tested the model on both pre-processed and raw data sets and compared it to 8 standard machine learning classifiers. After evaluation, they noticed that the suggested model outperforms machine learning classifiers by about 7% [13].

Abdelli et al. compared the performance of machine learning and deep learning in classifying Arabic texts in MSA and colloquial Algerian dialect using the SVM algorithm, which is one of the supervised machine learning algorithms, with the TFIDF technique for feature extraction. In contrast, they used LSTM, one of the DL algorithms with Continuous Bag of Words (CBOW) for Word Embedding. Two data sets were created to achieve the work. One for sentiment analysis included data collected by the author from Facebook using Facebook API, then categorized into positive and negative comments and added to three pre-existing data sets made by other researchers. After filtering the data and deleting the duplicated ones, the final dataset contained 49,864 elements, divided equally into positive and negative sentences. At the same time, the second dataset contained 1.4 gigabytes of Arabic Corpus merging from a pre-existing Arabic Corpus to be used in CBOW training. SVM outperformed LSTM in accuracy in this work, reaching 86% [14].

Ombabi et al. [15] combined five re-published data sets in many topics written in Arabic to be 15.100 reviews in total, divided into positive and negative classes equally. The authors proposed a Deep CNN-LSTM Arabic-SA model that works as follows:

After re-processing and normalization phase, they used FastText (Skip-gram and CBOW) as word embedding methods to be the input of CNN to extract features, then used fully connected LSTMs at last used SVM for classification and prediction. A confusion matrix measured the performance, and the accuracy was achieved at 90.75%. It is worth noting that the authors tested the efficiency of the proposed model using other classifiers (NB and KNN), but SVM was the classifier with the highest efficiency.

To ASA, Khalid Bolbol and Maghari applied KNN, LR, and DT classifiers on four pre-existing data sets in different sizes in Arabic, where ASTD's size was 399 KB, AJGT 160KB, ASTC 5.5 MB, and Arabic 100K Reviews 35.4 MB. Consequently, they pre-processed the data and extracted features by CBOW and TF-IDF techniques. The experiment showed that LR was the best accuracy [16].

The authors collected the dataset from Twitter about Covered-19 to classıfy ıt ınto posıtıve when there is fear or anxiety and negative when there is not a feeling of fear. They had balanced the collected dataset from social media for COVID-19 with pre-trained models CC.AR.300 and Arabic news. Then the authors pre-processed the dataset, and for feature extraction, using Word2Vec. The dataset was imbalanced, so they used SMOTENN on the training data to balance it. They utilized the SMOTE technique by synthesizing new minority samples and then merging or removing some samples from both classes (minority and majority); however, it uses k=3 nearest neighbors to locate samples in a misclassified dataset. They used nuSVC, LSVC, LRSV, SGD, and BNB. When applying word2vector with SMOTEEN on CC.AR.300 and in Arabic News, the performance was better than applying word embedding without SMOTEEN [17].

In this paper, the authors relied on Aspect Based Sentiment Analysis to analyze the sentiments of Arab customers in Saudi Arabia about the services of three Saudi telecom companies. Alshammari and Almansour collected 6182 tweets and then filtered and pre-processed them into 1096 tweets related to customer opinions about telecom companies. This dataset is divided into 80% for training and 20% for testing after being labeled manually. To extract features, they use unigram and bigram, then TF-IDF. At last, for sentiment analysis, the author compared the performance of SVM, LR, and RF with deep learning algorithms. The experiment showed that LR with unigram gave higher performance compared with the performance of the SVM, RF with unigram and bigram, and with the performance of LR with bigram. In contrast, the efficiency of deep learning performance with word embedding is better than the performance of machine learning classifiers and deep learning with POS, with an accuracy of 81% [18].

Elfaik and Nfaoui suggested using AraVec (CBOW) with BiLSTM with the attention model to apply it to three Arabic data sets: ASTD, ArTwitter, and Main AHS. The experiment showed that the performance of the proposed model outperformed the performance of several other algorithms used in previous studies [19].

In this paper, The proposed model architecture is DeepASA (Deep learning for Arabic Sentiment Analysis) which consists of an Input Layer required to transform the input into numerical representation using word embedding (FastText and word2vector). On the other hand, the structure of the DeepASA model has three main layers (input layer, hidden layer, and output layer). The hidden layer has two main neural networks (LSTM & GRU). For the dataset, the model experimented with many datasets (LABR, RES, HTL, PROD, ARTwitter, ASTD). DeepASA trained and tested with these datasets after being balanced for classification. They combined 3-different machine learning classifiers, and SVM, GNB, and LR produced the highest accuracy. Result: DeepASA scored the highest accuracy when trained and tested with the HTL dataset, and accuracy reached 94.32% [20].

In the case of this paper, Dataset: Arabic book reviews contain 63,257 reviews collected from Goodreads with ratings positive reviews reaching 42,832 and negative are 8224, and this is the balanced version of the dataset. For word embedding layers in proposal mode, they use Glove for word representation and FastText. The Neural network architecture (Model) uses the RNN, LSTM, and FastText models. LSTM model has the highest accuracy with 89.82 compared with the CNN model [21].

They created two different data sets in the Sudanese dialectical. SudSenti2 dataset was collected from Facebook and YouTube posts with 2,027 positive comments and 1,973 negative comments. At the same time, SudSenti3 data set was collected from Twitter with 2,523 positive tweets and 2,639 negative ones. after pre-processing the data, they used TF-IDF and the AraVec (word embedding). They offered a Sentiment

Convolutional Model (SCM) with Mean Max Average (MMA) pooling layer. They applied it to both created data sets and two pre-existed datasets in MSA (HARD dataset) and Saudi dialects (SSD dataset). The proposed model SCM+MMA showed the highest accuracy compared with traditional ML algorithms and CNN, RNN, and LSTM+CNN. The accuracy when the model applied on SudSenti2(2 classes) was 92.75%,on SSD(2 classes) was 85.55%, on HARD (2 classes) was 90.01%, while on SudSenti3(3classes) was 84.39%.[22].

Table 1. *Basic information in the literature*

Reviewed Paper	Domain, Dataset	Language	Algorithms	Features	Accuracy
[8]	Twitter, Microsoft Form	Saudi dialect	KNN, NB, LR, AVM, RF	TF-IDF	89% by SVM
[9]	Reviews from the Booking.com website	MSA, DA	LR, SVM, KNN, NB,DT,RF, GB, RC, MCL	Unigram, TF-IDF	95% by RC
[10]	Twitter	Sudanese dialect	KNN, NB, SVM	N-gram, TF-IDF	95% by SVM
[11]	YouTube comments	MSA, DA	KNN, Bernoulli NB, SVM with RBF (Radial Bases Function)	N-gram, TF-IDF	88% F-measure by SVM - RBF
[12]	Twitter	Saudi dialect	SVM	ANOVA based on F values	98% F1 score
[13]	Multi-Domain Reviews merged from: TDS, ASTD, SemEval, Social Media Posts, OCA, LABAR, LARGE, Health Services	MSA, DA	CNNs, KNN, SVM ,LR, DT, RF, BernoulliNB, Multinomial NB	character level (CNN), TF-IDF, unigram, bigram	94% by CNNs with character level.
[14]	Facebook comments	MSA, Algerian dialect	SVM, LSTM	TF-IDF, CBOW	86% by SVM and TF-IDF
[15]	Multi-Domain reviews	MSA, DA	CNN–LSTM, SVM, NB, KNN	Skip-gram, CBOW, CNN	SVM with Skipgram 90.75%
[16]	Multi-Domain reviews	MSA, DA	KNN, LR, and DT	CBOW , TF-IDF	93% by LR
[17]	Twitter, CC.AR.300, Arabic. news	MSA, DA	RF, SGD, SVM, BNB, and LRCV	CBOW and skip-gram	98.55% by NuSVM in CC.AR.300, 97.67% by NuSVM in Arabic. news
[18]	Twitter	Saudi dialect	SVM, LR, RF, Deep learning	unigram , bigram , TF-IDF	78% LR with unigram, Deep learning with word embedding 81%
[19]	ASTD, ArTwitter and Main AHS	MSA, DA	AraVec (CBOW) + BiLSTM + attention, Gaussian NB (GNB), RF, Nu-support vector (NuSYC), Stochastic gradient descent (SGD), (LR)	CBOW	By AraVec + BiLSTM + attention in ASTD 73%, in ArTwitter 83%, and in Main AHS, 89%
[20]	Multi-Domain reviews	MSA, DA	DeepASA LSTM & GRU	FastText Word2vector Doc2vector	With HTL dataset acc 94.32%
[21]	Arabic book reviews	MSA, DA	LSTM&CNN	Glove FastText	LSTM Model 89.82
[22]	SudSenti2, SudSenti3, SSD, HARD	MSA, Saudi dialect, Sudanese dialect	SCM+MMA, CNN, RNN, and CNN+LSTM, LR, RF, NB, and SVM	Word imbedding, TF-IDF, CNN	SCM+MMA model on SudSenti2 92.75%, on SudSenti3 84.39%, On SSD 85.55%, and on HARD 90.01%.

4. Discussion

In most papers, SVM shows high efficiency and good accuracy. Also, using CNN to extract features could increase the model's effectiveness. We notice that most of the papers show the phase of extracting the features from input after the pre-processing, and modern methods like FastText and Golve improve the accuracy up to 92% in most of the papers. On the other hand, extracting features using the traditional method does not reach an accuracy above 80%. For the proposal models trained and tested, the model LSTM &CNN with SVM classifier shows perfect results with an accuracy of 90.75%. However, the old model like KNN SVM, etc... does not reach high accuracy; most of the paper depends on traditional models to train that build on machine learning algorithms. Nevertheless, the model built on a deep learning algorithm shows the highest accuracy at 94.32 with the SVM classifier and Golve & FastText.

5. Conclusion

In this work, we presented a literature review of scientific papers in the field of ASA to help researchers know the best algorithms used in classification and in extracting features that have left a good impact in this field. In future work, we would like to use DL's algorithms to analyze Arab sentiments in Turkey.

Declaration of Interest

The authors declare that there is no conflict of interest.

References

[1] K. Jiang and X. Lu, "Natural Language Processing and Its Applications in Machine Translation: A Diachronic Review," *Proc. 2020 IEEE 3rd Int. Conf. Safe Prod. Informatiz. IICSPI 2020*, pp. 210–214, Nov. 2020, doi: 10.1109/IICSPI51290.2020.9332458.

[2] S. Al-Otaibi *et al.*, "Customer Satisfaction Measurement using Sentiment Analysis," *IJACSA) Int. J. Adv. Comput. Sci. Appl.*, vol. 9, no. 2, 2018, Accessed: Mar. 05, 2022. [Online]. Available: www.ijacsa.thesai.org.

[3] A. Das, K. S. Gunturi, A. Chandrasekhar, A. Padhi, and Q. Liu, "Automated Pipeline for Sentiment Analysis of Political Tweets," pp. 128–135, Jan. 2022, doi: 10.1109/ICDMW53433.2021.00022.

[4] K. Mishev, A. Gjorgjevikj, I. Vodenska, L. T. Chitkushev, and D. Trajanov, "Evaluation of Sentiment Analysis in Finance: From Lexicons to Transformers," *IEEE Access*, vol. 8, pp. 131662–131682, 2020, doi: 10.1109/ACCESS.2020.3009626.

[5] R. Duwairi and F. Abushaqra, "Syntactic- and morphology-based text augmentation framework for Arabic sentiment analysis," *PeerJ Comput. Sci.*, vol. 7, pp. 1–25, 2021, doi: 10.7717/PEERJ-CS.469/.

[6] S. V. Pandey and A. V. Deorankar, "A Study of Sentiment Analysis Task and It's Challenges," *Proc. 2019 3rd IEEE Int. Conf. Electr. Comput. Commun. Technol. ICECCT 2019*, Feb. 2019, doi: 10.1109/ICECCT.2019.8869160.

[7] Y. Zahidi, Y. E. L. Younoussi, and Y. Al-Amrani, "Arabic Sentiment Analysis Problems and Challenges," *Proc. - 10th Int. Conf. Virtual Campus, JICV 2020*, Dec. 2020, doi: 10.1109/JICV51605.2020.9375650.

[8] L. M. Alharbi and A. M. Qamar, "Arabic Sentiment Analysis of Eateries' Reviews: Qassim region Case study," *Proc. - 2021 IEEE 4th Natl. Comput. Coll. Conf. NCCC 2021*, Mar. 2021, doi: 10.1109/NCCC49330.2021.9428788.

[9] A. A. Sayed, E. Elgeldawi, A. M. Zaki, and A. R. Galal, "Sentiment Analysis for Arabic Reviews using Machine Learning Classification Algorithms," *Proc. 2020 Int. Conf. Innov. Trends Commun. Comput. Eng. ITCE 2020*, pp. 56–63, Feb. 2020, doi: 10.1109/ITCE48509.2020.9047822.

[10] S. Abuuznien, Z. Abdelmohsin, E. Abdu, and I. Amin, "Sentiment Analysis for Sudanese Arabic Dialect Using comparative Supervised Learning approach," *Proc. 2020 Int. Conf. Comput. Control. Electr. Electron. Eng. ICCCEEE 2020*, Feb. 2021, doi: 10.1109/ICCCEEE49695.2021.9429560.

[11] A. K. Al-Tamimi, A. Shatnawi, and E. Bani-Issa, "Arabic sentiment analysis of YouTube comments," *2017 IEEE Jordan Conf. Appl. Electr. Eng. Comput. Technol. AEECT 2017*, vol. 2018-January, pp. 1–6, Jul. 2017, doi: 10.1109/AEECT.2017.8257766.

[12] M. Alassaf and A. M. Qamar, "Aspect-Based Sentiment Analysis of Arabic Tweets in the Education Sector Using a Hybrid Feature Selection Method," *Proc. 2020 14th Int. Conf. Innov. Inf. Technol. IIT 2020*, pp. 178–185, Nov. 2020, doi: 10.1109/IIT50501.2020.9299026.

[13] E. Omara, M. Mosa, and N. Ismail, "Deep Convolutional Network for Arabic Sentiment Analysis," *2018 Proc. Japan-Africa Conf. Electron. Commun. Comput. JAC-ECC 2018*, pp. 155–159, Apr. 2019, doi: 10.1109/JEC-ECC.2018.8679558.

[14] A. Abdelli, F. Guerrouf, O. Tibermacine, and B. Abdelli, "Sentiment Analysis of Arabic Algerian Dialect Using a Supervised Method," *Proc. - 2019 Int. Conf. Intell. Syst. Adv. Comput. Sci. ISACS 2019*, Dec. 2019, doi: 10.1109/ISACS48493.2019.9068897.

[15] A. H. Ombabi, W. Ouarda, and A. M. Alimi, "Deep learning CNN–LSTM framework for Arabic sentiment analysis using textual information shared in social networks," *Soc. Netw. Anal. Min. 2020 101*, vol. 10, no. 1, pp. 1–13, Jul. 2020, doi: 10.1007/S13278-020-00668-1.

[16] N. Khalid Bolbol and A. Y. Maghari, "Sentiment analysis of arabic tweets using supervised machine learning," *Proc. - 2020 Int. Conf. Promis. Electron. Technol. ICPET 2020*, pp. 89–93, Dec. 2020, doi: 10.1109/ICPET51420.2020.00025.

[17] W. Al-Sorori *et al.*, "Arabic Sentiment Analysis towards Feelings among Covid-19 Outbreak Using Single and Ensemble Classifiers," pp. 1–6, Nov. 2021, doi: 10.1109/ITSS-IOE53029.2021.9615256.

[18] N. F. Alshammari and A. A. Almansour, "Aspect-based Sentiment Analysis for Arabic Content in Social Media," *2nd Int. Conf. Electr. Commun. Comput. Eng. ICECCE 2020*, Jun. 2020, doi: 10.1109/ICECCE49384.2020.9179327.

[19] H. Elfaik and E. H. Nfaoui, "Deep attentional bidirectional LSTM for arabic sentiment analysis in Twitter," *2021 1st Int. Conf. Emerg. Smart Technol. Appl. eSmarTA 2021*, Aug. 2021, doi: 10.1109/ESMARTA52612.2021.9515751.

[20] A. Alharbi, M. Kalkatawi, and M. Taileb, “Arabic Sentiment Analysis Using Deep Learning and Ensemble Methods,” *Arab. J. Sci. Eng. 2021 469*, vol. 46, no. 9, pp. 8913–8923, May 2021, doi: 10.1007/S13369-021-05475-0.

[21] M. Fawzy, M. W. Fakhr, and M. A. Rizka, “Word Embeddings and Neural Network Architectures for Arabic Sentiment Analysis,” *16th Int. Comput. Eng. Conf. ICENCO 2020*, pp. 92–96, Dec. 2020, doi: 10.1109/ICENCO49778.2020.9357377.

[22] M. Mhamed, R. Sutcliffe, X. Sun, J. Feng, E. Almekhlafi, and E. A. Retta, “A Deep CNN Architecture with Novel Pooling Layer Applied to Two Sudanese Arabic Sentiment Datasets,” Jan. 2022, Accessed: Jun. 03, 2022. [Online]. Available: http://arxiv.org/abs/2201.12664.