Arabic Sentiment Analysis for Twitter Data: A Systematic Literature Review

Tahani Alqurashi
College of Computer and Information Systems, Information System Department, Umm Al-Qura University, Saudi Arabia
tmqurashi@uqu.edu.sa
(corresponding author)

Received: 5 January 2023 | Revised: 22 January 2023 and 24 January 2023 | Accepted: 26 January 2023

ABSTRACT

Social media platforms have a huge impact on our daily lives. They have succeeded in attracting many people to spend time communicating and expressing themselves. Twitter is a social media platform that could be considered as a source of public opinion about products, services, and events. Sentiment analysis is the art of studying public feelings about certain topics, which may be positive, negative, or neutral. This paper provides a systematic review of Arabic tweet sentiment analysis on papers published from 2012 to 2021 in digital libraries including IEEE Explorer, Science Direct, Springer Link, and Google Scholar. The main aim of this systematic review is to investigate the trends in the topics reported and to highlight potential new research lines. To achieve that, three main stages were implemented: planning, conducting, and reporting the review. Our findings suggest the need for an open-source large Arabic tweet dataset that can be used by researchers. Also, it was found that researchers have used various classification techniques, which led to different results.

Keywords-arabic sentiment analysis; systematic review; social media; twitter

I. INTRODUCTION

Social media are considered a major source of news, marketing, and advertisements. These platforms have huge numbers of users, which are increasing daily. The number of global monthly active users of Twitter recently averaged to 330 million people [1]. Users express their feelings and opinions using short messages called tweets and they can refer tweets to other users, called followers. Therefore, the content of these platforms may be evaluated to extract insights from these data.

Sentiment Analysis (SA), also known as opinion mining, is one of the natural language processing fields. SA focuses on analyzing text to extract people's opinions and emotions and identify these as positive, negative, or neutral [2]. Many researchers have directed their attention to this area and while most research has focused on English, Chinese, and other Indo-European languages, few studies have addressed SA in morphologically rich languages, such as Arabic. The number of Arabic texts existing on the internet has seen a significant increase. According to Internet World Statistics, Arabic is the fourth most commonly used language on the Internet, after English, Chinese, and Spanish, reaching 5.2% of all internet users [3]. Due to the exponential growth in Arabic internet users and Arabic online content, Arabic SA (ASA) has gained the attention of many researchers over the last decade [4].

In this paper, we systematically review the literature in the field of Arabic tweet SA.

II. OVERVIEW OF THE SYSTEMATIC REVIEW METHOD

In this review, the systematic literature review methodology proposed in [5] was followed. It consists of 3 main stages: planning, conducting, and reporting the review. The main aim of this review is to perform a comprehensive review of ASA for Twitter data and cover the methodologies and approaches of data processing, SA techniques and approaches that have been used in the literature, and the challenges ASA faces. Figure 1 illustrates the steps of the systematic literature review methodology, which are explained in more detail in the following sections.

A. *Research Questions*

In the planning phase, the research questions are specified. In this study, the research contributions are highlighted through answering the following Research Questions (RQs):

- RQ1: What are the different machine learning techniques that have been applied/proposed for ASA?
- RQ2: What data pre-processing techniques have been used for Arabic tweets in ASA research?

B. *Search Strategy*

After determining the research questions, the search terms and data resources are specified. First, to identify the search terms, the research questions were analyzed and the following search strings were developed: "Arabic", "Arabic text", "sentiment analysis", "Twitter", and "tweets". Moreover, the Boolean operators "OR" and "AND" were used to search for all possible combinations of these terms. We covered only ASA studies using Twitter data as the application.

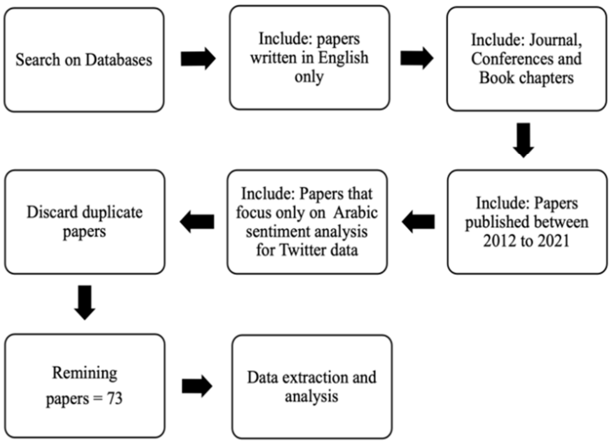

Fig. 1. The steps of the systematic literature review methodology.

Secondly, a search was carried out on the following six academic databases: IEEE Explore, ACM, Springer, ScienceDirect/Elsevier, Google Scholar, and Wiley. These were chosen as they are the top databases in the computer science field. The titles, abstracts, and keywords of all indexed papers were searched using the search terms we developed, and the search was conducted on studies from 2012 up to 2021.

TABLE I. SEARCH RESULTS

Database	Relevant search results
IEEE Explore	38
ACM	4
ScienceDirect/Elsevier	9
Springer Link	5
Google Scholar	16
Wiley	1
Total	73

C. Inclusion and Exclusion Criteria

In this stage, inclusion and exclusion criteria were delineated to thoroughly assess the relevance of the potential primary studies. These were:

- Include papers written in English only.
- Include papers published in journals, conference papers, and book chapters.
- Include papers published from 2012 to 2021.
- Include papers that focus on Arabic text analysis on the Twitter platform. Exclude papers that focus on all other platforms.
- Exclude duplicate papers.
- Exclude secondary studies (i.e. literature reviews).
- Exclude papers that are not peer-reviewed, such as technical reports and theses.

D. Data Extraction

After collecting the data from the above-mentioned databases and extracting the relevant papers based on our inclusion and exclusion criteria, 73 papers were selected for analysis. Table I shows the search results. The papers were classified based on the type and the year of publication. From 2012 to 2021, a gradual increase in the numbers of conference papers and published journal articles was noted, as shown in Figure 2. The number of publications peaked in 2019, with 19 identified publications in total. The average number of publications is around 9 studies per year over the past 10 years. The most common form of publication was conference papers (42), followed by journal articles (30), and 1 book chapter, as shown in Figure 3.

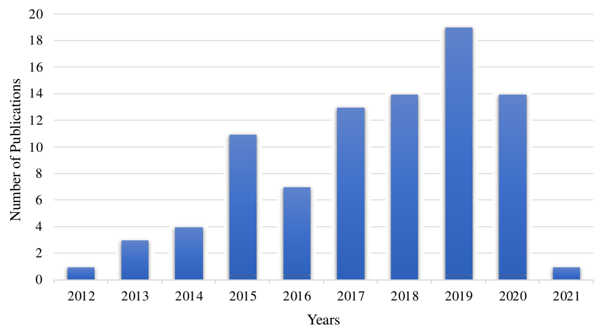

Fig. 2. Number of publications per year.

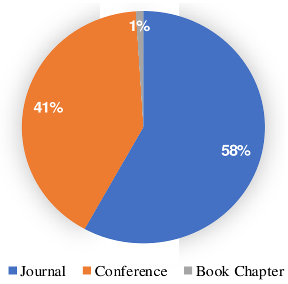

Fig. 3. Type of publications.

III. PRIMARY STUDIES AND DISCUSSION

SA is a set of processes that are applied widely in computer science studies to analyze and examine semantics, words, and tweet syntax to determine the emotions in the text [6]. The main purpose of SA using Twitter data is to classify tweets into three polarities (positive, negative, or natural) based on the statements or words contained in that tweet [7]. In this section, the main finding of this systematic review regarding ASA using the Twitter platform are presented and discussed.

A. Arabic Tweet Sentiment Analysis Approaches

We found that ASA approaches can be classified into three main categories: corpus-based, lexicon-based, and hybrid-based. Table II summarizes the reviewed papers based on their categories, The following subsections discuss each of these separately.

1) The Corpus-based Approach

In this approach, the main aim was to collect and use the available Twitter dataset to build machine learning models in order to determine the sentiments of each tweet. Our findings indicate that the most commonly used methods are the supervised learning approach, and in particular Naive Bayes (NB), Support Vector Machine (SVM), Decision Tree (DT), K-Nearest Neighbor (KNN), Ensemble classifiers (bagging and voting), and deep learning [2, 8-11]. Moreover, a few studies used an unsupervised learning approach [12, 13].

In the supervised learning approach, authors in [8] used SVM, multinomial NB and rule-based classification algorithms to classify sentiments based on their collected dataset of 134,194 Arabic tweets, which had been labelled automatically using emojis. They concluded that SVM outperformed the other algorithms used, reaching an accuracy of 75.7%. Authors in [9] investigated the state-of-art English SA techniques and used them for solving the problem of ASA through a Machine Translation (MT)-based approach. They first collected Arabic tweets (a total of 937 tweets), translated them into English, and then applied Stanford sentiment classifiers to classify each tweet, either as positive or negative sentiment. By comparing this method with lexicon-based techniques, they found that their MT sentiment approach outperformed the other approaches. Authors in [2] utilized machine learning algorithms and neural networks to perform SA on health services. They collected Twitter data using trending hashtags about health services, and their results show that the accuracy of the SVM approach outperformed the other algorithms, including CNN and DNN. Authors in [10] collected Arabic tweets related to political events in 2018. They analyzed their dataset using machine learning and data mining algorithms such as NB, SVM, DT, KNN, and ensemble classifiers. They observed that ensemble, and in particular the voting classifier, outperformed the other techniques. Authors in [14] examined tweets on women's issues in Saudi Arabia. They collected their dataset from popular Twitter hashtags, and classified them as positive, negative, and neutral. The NB algorithm was used to automate classification. Authors in [15] built a predictive model and applied it to data gathered from the Gulf region to predict whether a user's tweet indicated a depressed or non-depressed user. They applied the common machine learning algorithms, RF, NB, AdaBoost1 and linear SVM, and concluded that linear SVM outperformed the other classifiers reaching an accuracy of 87.5%. Authors in [16] aimed to investigate sentiments from tweets that were collected from popular hashtags about women's problems in Saudi Arabia. The data was pre-processed and classified, and their results indicated that the NB algorithm achieved greater accuracy than the other classifiers available in Weka.

In a deep learning approach, authors in [17] proposed a Multi-Channel Embedding Convolutional Neural Network (MCE-CNN) for Arabic sentiment classification and thoroughly assessed it using balanced and imbalanced datasets. The MCE-CNN was trained on different sentiment features, including word and character n-grams. Their proposed model achieved high accuracy. Authors in [18] proposed a sentiment and emotion system for regression and classification tasks using the CNN and LSTM algorithms. They implemented different feature extraction techniques, a word embedding model (Aravec), a document embedding model (Doc2vec), and a set of semantic features such as Deepemoji and unsupervised sentiment Nerous. Their proposed system outperformed the SVM unigrams baseline model's performance applied to the SemEval2018 dataset [19]. Authors in [20] developed a new architecture based on CNNs and RNNs for handwritten Arabic word understanding and classification. The most powerful technique for analyzing Arabic tweets and social media stata is the CNN technique. The use of a sub-word-level RNN module and a character-level CNN module helps in gaining a better understanding of handwritten text. This technique is efficient, especially with a short text written in an uncontrolled language format. An approach to tweet-based word correction that uses Arabic text classification to find abusive accounts on Twitter was proposed in [21]. The classification of Arabic tweets without stemming was investigated and compared to classification using stemming. This approach proved better than other popular approaches to Arabic word correction on Twitter. Authors in [22] proposed the classification of Arabic tweets as positive or negative using an improved algorithm. The approach consists of two phases. In the first phase the dataset was prepared by normalization, while the second phase involved classifying tweets using a DMNB approach. The experimental results showed that this improved approach reached an accuracy level up to 87.5%.

In an unsupervised learning approach, authors in [12] used sentiment classification of people's opinions during the Covid 19 pandemic using k-means and mini-batch k-means algorithms on Arabic and English datasets. They observed that k-means classification takes longer than mini-batch k-means classification. Authors in [13] aimed to assess the integration of the similarity functions with pre-processing methods for clustering tweets. K-means clustering algorithms were used to cluster Arabic tweets into two clusters: positive and negative. The results showed that the stemming (pre-processing method) with the Kullback-Leibler divergence function is more effective than other competitive pre-processing techniques.

Other studies compared different approaches. Authors in [23] compared the corpus-based and lexicon-based approaches. For the first approach, they collected 2000 Arabic tweets and used SVM, NB, KNN, and DT algorithms. For the lexicon approach they manually built a lexicon of 3479 words. They concluded that the SVM approach outperformed the others. Authors in [24] applied three types of classification methods (supervised, unsupervised, and hybrid learning) on tweets collected randomly from three different domains (sports, politics, and social). The results indicate that hybrid learning is better than the supervised or the unsupervised approaches in terms of classification accuracy.

2) The Lexicon-based Approach

This approach, considered to be unsupervised learning, aims to build a sentiment lexicon of terms and then compute the sentiment of a certain text based on the sentiment values of the terms composing it. Our findings indicate that there are two different lexicon-based approaches used in ASA: the Double Polarity (DP) [25] and the Simple Polarity (PL) [26] approaches. The latter is based on counting both the positive and negative words in each sentence, while the DP approach is based on the frequencies of positive and negative words in the sentence. In [26], a PL lexicon was built manually, consisting of 3982 adjectives from two Arabic datasets that were classified as positive, negative, and neutral. The authors applied PL to classify individual tweets on the basis of including positive, negative or neutral adjectives. They called their system SAMAR and they concluded that using PL improves the accuracy of ASA. Authors in [27] proposed a Weighted Lexicon-Based Algorithm (WLBA) for the Saudi dialect. In their approach, the weight of each word was calculated based on the corpus, not upon the lexicon. Moreover, they compared their proposed algorithm with the DP and PL algorithms, and concluded that WLBA outperformed the DP approach, but not the PL approach.

The lexicon-based approach depends on the creation of a lexicon. We found that there are two different methods of lexicon creation: manual and automatic. The manual method is the most popular and it provides more accurate results, but is domain-independent [27-29].

TABLE II. REVIEWED PAPERS CLASSIFIED BASED ON THE APPROACH USED

Approach	Paper	Dataset size	Algorithms	Features	Language
Corpus-based	[8]	134194 tweets	SVM, MNB and rule-based	TF-IDF	MSA, dialects
	[10]	14419 tweets	SVM, NB, NT, KNN, Ensemble (bagging, voting)	TF-IDF	Unspecified
	[14]	9096	SVM, NB	BOW, bi-gram, and tri-gram	Dialects
	[15]	6122 tweets from 89 users	RF, SVM, NB, Adaboost and liner SVM	Unigrams, negative features	MSA, Gulf dialect
	[17]	ASTD dataset, 10000 tweets	MCE-CNN	Word and character n-grams	MSA, dialects
	[20]	ASTD	Stacked model with CNN and gated recurrent unit	Character level model	MSA
	[37]		NB, SVM, multinomial logistic regression, KNN		Dialect
	[64]	ASTD (3,315 tweets)	SVM, recursive neural tensor network	Character n-gram model [3,5], Word n-grams [1,4], count of negated words and positive and negative words based on lexicons from [65-67]	Egyptian dialect
	[68]	2591 tweets	SVM, NB, KNN		MSA and dialects
	[2]	2026 tweets	SVM, NB, LR, CNN, DNN	TF-IDF, word frequency, word2Vec	MSA
	[70]		SVM		MSA, dialects
	[71]	17748 tweets	SVM	TF-IDF	MSA, dialects
	[72]	SemEval, AraSenTi, ASTD	Deep learning, SVM	Word embedding (AreVec), set of related lexicon features	MSA, dialects
	[73]	2.3M tweets on news	Semi-supervised technique		MSA, dialects
	[9]	937 tweets	Machine translation approach, Stanford sentiment classifier		MSA
Lexicon-based	[69]	22550	SVM, NB	Binary model	Jordanian dialects
	[26]	3015 tweets, 2798 tweets	Polarity lexicon of size 3982		MSA, dialects
	[28]	2000 tweets on politics and arts	Manual lexicon of size 4815 words		MSA, Jordanian dialects
	[9]	937 tweets		Machine translation approach, Stanford sentiment classifier	MSA
	[52]		SVM		MSA, dialects
	[73]	5400 tweets and emoji data			MSA, Saudi dialect
	[74]	6000 tweets	Rule-based algorithm	Semantic and lexicons features	MSA, Tunisian dialect
Hybrid-based	[75]	Merging of two available lexicons (MPQA and ArabSenti) with manually collected lexica	SVM, LIBSVM	Stems, Twitter language independent, semantic features	Egyptian dialects
	[53]	4800 tweets	Semantic orientation, SVM, NB	unigram, bi-gram, tri-gram	Egyptian dialect
	[30]	Lexicon of 5376 words	RF, SVM, Max-Ent, bagging, boosting, ANN, DT, NB	POS and lexicon features	MSA, dialect
	[31]	1500 tweets and lexicon of 452 words	Bagging and boosting		MSA, dialect
	[18]	SemEval2018	Regression and classification tasks, CNN, LSTM	Word embedding (AraVec), document embedding (doc2Vec) and a set of semantic features	MSA, Egyptian and Gulf dialects

3) The Hybrid-based Approach

Authors in [30, 31] worked on SA for Arabic tweets using a hybrid approach. In [30], a combination of the lexicon and corpus techniques for ASA was proposed. The authors used a lexicon to replace words in sentences with their sentiment labels to enable classifier algorithms to consider rare and important words in the corpus. They concluded that this hybrid approach outperformed the corpus-based approach, reaching an accuracy of 96.3%. The hybrid approach in [31] comprised machine learning algorithms, the bagging and boosting approaches. The authors started by building a unified dataset of text and audio which was later converted to text, to analyze sentiments. However, some of the audio analysis was not accurate, which had a negative effect on this prediction approach because some audios contained laughing and yelling. Authors in [32] proposed an incremental learning system for sentiment classification. Their system was able to update a lexicon with any upcoming changes. In fact, this system integrated the machine learning and lexicon approaches to improve the accuracy of the proposed system, and it has been tested with different datasets.

B. Tools Used to Collect Arabic Tweets

As we only focus on the Twitter platform, we found various techniques have been used to collect Arabic tweets. Most studies used Twitter's official Application Programming Interface (API) to collect Arabic tweets [2, 33-39]. Other studies used trending hashtags to extract data manually [8, 40]. Authors in [16] used Keyhole and Netlytic tools to collect their data from Twitter. Authors in [41] developed a Python script connecting with Twitter's official API to collect their data. Authors in [42] developed a Twitter data grabber tool, a small desktop application using C# to connect to the Twitter API to collect their data. Authors in [10, 15] used a Java library called Twitter4J, which can be connected to the Twitter API to collect Twitter data. Authors in [43] used a tool called Archivist4 to collect tweets using hashtags. In fact, Archivist4 can be used to archive and analyze visual tweets using hashtags, usernames, Boolean, and complex search terms.

C. Data Pre-Processing

Data pre-processing is a technique used to prepare data for sentiment classification [78, 79]. The preparation involves cleaning, formatting and sorting the tweets that have been collected from Twitter to be saved in a dataset ready for analysis [45]. Data pre-processing for Arabic tweets includes several steps to clean the data, such as tokenization, stop-word removal, text cleaning, Part Of Speech (POS) tagging, normalization, and stemming. Table III shows the most commonly used methods for data pre-processing found in the papers we reviewed. We note that normalization, text cleaning and tokenization were the most frequently used techniques in these papers. These techniques are described below.

TABLE III. DATA PRE-PROCESSING METHODS USED IN THE REVIEWED PAPERS

Reference	Tokenization	Stop word removal	Text cleaning	POS tagging	Normalization	Stemming
[23]	∅	✓	✓	∅	✓	✓
[10]	✓	✓	∅	∅	∅	✓
[51-53]	∅	✓	✓	∅	✓	✓
[15]	✓	✓	✓	∅	∅	✓
[44, 46]	✓	✓	✓	∅	✓	✓
[54]	✓	∅	∅	∅	✓	∅
[55]	✓	∅	✓	∅	✓	∅
[56]	✓	∅	∅	✓	✓	∅
[42]	✓	✓	∅	∅	✓	✓
[45]	✓	∅	✓	∅	✓	✓
[57]	✓	✓	✓	∅	∅	✓
[8]	∅	✓	✓	∅	✓	∅
[58]	✓	✓	∅	∅	✓	∅
[47, 59]	✓	∅	✓	✓	✓	∅
[30]	∅	∅	✓	∅	✓	∅
[60]	∅	✓	∅	∅	∅	✓
[66]	∅	∅	✓	∅	✓	✓
[61]	✓	∅	✓	∅	✓	∅
[62]	∅	✓	✓	∅	✓	✓
[32]	✓	∅	✓	∅	✓	∅
[63]	∅	∅	∅	∅	✓	✓
[33]	∅	✓	✓	✓	✓	∅

- Tokenization is the process of the segmentation of a stream of text to segments consisting of words or phrases. Each segment is called a token and each token has a meaning and can be used for the later stages of SA [46].
- Stop-word removal involves removing words that have no meaning for polarity classification [42], such as (من, على).
- Text clearing: Arabic tweets may have inconsistent text and can be noisy or incomplete. Therefore, data cleaning tends to remove noise, complete missing values, or correct inconsistent state data [33]. This also including removing Arabic articles [8], such as (ال، بال، كال).
- POS tagging involves mapping words to their tags, such as verbs, nouns and adjectives. There are different types of machine tools for POS tagging [47].

- Normalization: Most reviewed papers applied normalization to Arabic tweets, which includes removing repetitive letters or normalizing similar letters to the same letter [42], such as the letters (أ-آ) which is normalized to (ا).

- Stemming is the process of transforming a word to its base forms, while the meaning of the words is preserved. For example, in Arabic the word (يكتبو□ ,يكاتبو□) is replaced with (كتب). As the Arabic language has a complicated morphological structure, stemming is considered a very difficult task. Our findings indicate that there are two popular steaming methods used in ASA: light10 stemming [48] and Khoja's stemmer [49].

IV. CHALLENGES IN ARABIC SENTIMENT ANALYSIS

The Arabic language is one of the most widely spoken languages in the world. It is the most frequently spoken and written language in the Arab world, especially in the Middle East and North African regions [77]. The Arabic language has 26 letters and is written from right to left. It uses diacritical marks that denote correct pronunciation of a word. In addition, diacritical marks are used to distinguish words that have the same letters but different meanings. Moreover, the Arabic language has three forms: Classical Arabic, which is seen in religious and very formal texts, Modern Standard Arabic (MSA), which is associated with modern news media [77], and dialect or informal Arabic, which is the Arabic spoken with different local accents across the Middle East and North African countries, and has no particular standards [6]. According to [50], Saudi Twitter covers 90% of dialects, compared to MSA. As a result, it is very challenging for researchers to construct models of Arabic text classifiers to use for SA. In terms of translation, when translating Arabic to English, good results may be obtained with MSA, but Arabic dialects are difficult to translate because their meaning is purely context-related [27].

V. CONCLUSION AND FUTURE WORK

The Twitter platform is considered a rich source for sentiment analysis. People use social media platforms to express their opinions of products, services, or political events. In this paper, we have provided a systematic review of Arabic tweet sentiment analysis. We have reviewed the main techniques that can be used to prepare the data for sentiment analysis. Moreover, we have presented some of the major tools used for sentiment analysis of Arabic tweets. However, this review has some limitations: it was carried out using six academic databases and we only included papers that were written in English. Also, we did not include a secondary studies paper.

We note from the papers we reviewed that researchers have used various classification techniques, which has led to different results due to the lack of experiments applied to a standardized dataset. In the future, researchers should narrow their research domains and focus on Arabic dialects. They should also investigate complex classification techniques such as ensemble machine learning techniques and deep learning models. Lastly, we noticed that there is a lack of open-source large Arabic tweet datasets that can be used by researchers. Therefore, more work needs to be done in building large open-source databases of Arabic tweets.

REFERENCES

[1] "Twitter global mDAU 2022," *Statista*, Nov. 11, 2022. https://www.statista.com/statistics/970920/monetizable-daily-active-twitter-users-worldwide/.

[2] A. M. Alayba, V. Palade, M. England, and R. Iqbal, "Arabic language sentiment analysis on health services," in *1st International Workshop on Arabic Script Analysis and Recognition*, Nancy, France, Apr. 2017, pp. 114–118, https://doi.org/10.1109/ASAR.2017.8067771.

[3] "World Arabic Language Day 2020: focus on Arabic Language Academies | UNESCO." https://www.unesco.org/en/articles/world-arabic-language-day-2020-focus-arabic-language-academies (accessed Jan. 28, 2023).

[4] N. Boudad, R. Faizi, R. Oulad Haj Thami, and R. Chiheb, "Sentiment analysis in Arabic: A review of the literature," *Ain Shams Engineering Journal*, vol. 9, no. 4, pp. 2479–2490, Dec. 2018, https://doi.org/10.1016/j.asej.2017.04.007.

[5] B. Kitchenham, "Procedures for Performing Systematic Reviews," Keele University, Keele, UK, Technical Report TR/SE-0401, 2004.

[6] M. Salameh, S. Mohammad, and S. Kiritchenko, "Sentiment after Translation: A Case-Study on Arabic Social Media Posts," in *Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Denver, CO, USA, Jun. 2015, pp. 767–777, https://doi.org/10.3115/v1/N15-1078.

[7] D. Kurniawati, E. Prayitno, D. F. Sari, and S. N. Putra, "Sentiment Analysis of Twitter Use on Policy Institution Services using Naive Bayes Classifier Method," *Journal of International Conference Proceedings*, vol. 2, no. 1, Apr. 2019, Art. no. 33, https://doi.org/10.32535/jicp.v2i1.409.

[8] N. Boudad, R. Faizi, R. Oulad Haj Thami, and R. Chiheb, "Sentiment analysis in Arabic: A review of the literature," *Ain Shams Engineering Journal*, vol. 9, no. 4, pp. 2479–2490, Dec. 2018, https://doi.org/10.1016/j.asej.2017.04.007.

[9] E. Refaee and V. Rieser, "Benchmarking Machine Translated Sentiment Analysis for Arabic Tweets," in *Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Denver, CO, USA, Jun. 2015, pp. 71–78.

[10] H. M. Najadat, A. A. Alzu'bi, F. Shatnawi, S. Rawashdeh, and W. Eyadat, "Analyzing Social Media Opinions Using Data Analytics," in *11th International Conference on Information and Communication Systems*, Irbid, Jordan, Apr. 2020, pp. 266–271, https://doi.org/10.1109/ICICS49469.2020.239497.

[11] R. M. Alahmary, H. Z. Al-Dossari, and A. Z. Emam, "Sentiment Analysis of Saudi Dialect Using Deep Learning Techniques," in *International Conference on Electronics, Information, and Communication*, Auckland, New Zealand, Jan. 2019, pp. 1–6, https://doi.org/10.23919/ELINFOCOM.2019.8706408.

[12] M. A. Alanezi and N. M. Hewahi, "Tweets Sentiment Analysis During COVID-19 Pandemic," in *International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy*, Sakheer, Bahrain, Oct. 2020, pp. 1–6, https://doi.org/10.1109/ICDABI51230.2020.9325679.

[13] D. Abuaiadah, D. Rajendran, and M. Jarrar, "Clustering Arabic Tweets for Sentiment Analysis," in *14th International Conference on Computer Systems and Applications*, Hammamet, Tunisia, Nov. 2017, pp. 449–456, https://doi.org/10.1109/AICCSA.2017.162.

[14] G. Alwakid, T. Osman, and T. Hughes-Roberts, "Challenges in Sentiment Analysis for Arabic Social Networks," *Procedia Computer Science*, vol. 117, pp. 89–100, Jan. 2017, https://doi.org/10.1016/j.procs.2017.10.097.

[15] S. Almouzini, M. khemakhem, and A. Alageel, "Detecting Arabic Depressed Users from Twitter Data," *Procedia Computer Science*, vol. 163, pp. 257–265, Jan. 2019, https://doi.org/10.1016/j.procs.2019.12.107.

[16] E. Alyami, S. Matwin, T. Elmasri, and H. Ali-Hassan, "Mining The Saudi Twittersphere for Expressive Support Towards Women," in *Fourth World Conference on Smart Trends in Systems, Security and Sustainability*, London, UK, Jul. 2020, pp. 301–305, https://doi.org/10.1109/WorldS450073.2020.9210367.

[17] A. Dahou, S. Xiong, J. Zhou, and M. A. Elaziz, "Multi-Channel Embedding Convolutional Neural Network Model for Arabic Sentiment Classification," *ACM Transactions on Asian and Low-Resource Language Information Processing*, vol. 18, no. 4, pp. 1–23, Feb. 2019, https://doi.org/10.1145/3314941.

[18] M. Abdullah, M. Hadzikadicy, and S. Shaikhz, "SEDAT: Sentiment and Emotion Detection in Arabic Text Using CNN-LSTM Deep Learning," in *17th IEEE International Conference on Machine Learning and Applications*, Orlando, FL, USA, Dec. 2018, pp. 835–840, https://doi.org/10.1109/ICMLA.2018.00134.

[19] S. Mohammad, F. Bravo-Marquez, M. Salameh, and S. Kiritchenko, "SemEval-2018 Task 1: Affect in Tweets," in *12th International Workshop on Semantic Evaluation*, New Orleans, LA, USA, Jun. 2018, pp. 1–17, https://doi.org/10.18653/v1/S18-1001.

[20] M. Beseiso and H. Elmousalami, "Subword Attentive Model for Arabic Sentiment Analysis: A Deep Learning Approach," *ACM Transactions on Asian and Low-Resource Language Information Processing*, vol. 19, no. 2, pp. 1–17, Oct. 2020, https://doi.org/10.1145/3360016.

[21] E. A. Abozinadah and J. H. Jones, "Improved Micro-Blog Classification for Detecting Abusive Arabic Twitter Accounts," *International Journal of Data Mining & Knowledge Management Process*, vol. 6, no. 6, pp. 17–28, Nov. 2016, https://doi.org/10.5121/ijdkp.2016.6602.

[22] H. AlSalman, "An Improved Approach for Sentiment Analysis of Arabic Tweets in Twitter Social Media," in *3rd International Conference on Computer Applications & Information Security*, Riyadh, Saudi Arabia, Mar. 2020, pp. 1–4, https://doi.org/10.1109/ICCAIS48893.2020.9096850.

[23] N. A. Abdulla, N. A. Ahmed, M. A. Shehab, and M. Al-Ayyoub, "Arabic sentiment analysis: Lexicon-based and corpus-based," in *IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies*, Amman, Jordan, Dec. 2013, pp. 1–6, https://doi.org/10.1109/AEECT.2013.6716448.

[24] S. Alhumoud, T. Albuhairi, and W. Alohaideb, "Hybrid sentiment analyser for Arabic tweets using R," in *7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*, Lisbon, Portugal, Nov. 2015, vol. 1, pp. 417–424.

[25] S. R. El-Beltagy and A. Ali, "Open issues in the sentiment analysis of Arabic social media: A case study," in *9th International Conference on Innovations in Information Technology*, Al Ain, United Arab Emirates, Mar. 2013, pp. 215–220, https://doi.org/10.1109/Innovations.2013.6544421.

[26] M. Abdul-Mageed, M. Diab, and S. Kubler, "SAMAR: Subjectivity and sentiment analysis for Arabic social media," *Computer Speech & Language*, vol. 28, no. 1, pp. 20–37, Jan. 2014, https://doi.org/10.1016/j.csl.2013.03.001.

[27] A. Assiri, A. Emam, and H. Al-Dossari, "Towards enhancement of a lexicon-based approach for Saudi dialect sentiment analysis," *Journal of Information Science*, vol. 44, no. 2, pp. 184–202, Apr. 2018, https://doi.org/10.1177/0165551516688143.

[28] N. A. Abdulla, N. A. Ahmed, M. A. Shehab, M. Al-Ayyoub, M. N. Al-Kabi, and S. Al-rifai, "Towards Improving the Lexicon-Based Approach for Arabic Sentiment Analysis," *International Journal of Information Technology and Web Engineering*, vol. 9, no. 3, pp. 55–71, Jul. 2014, https://doi.org/10.4018/ijitwe.2014070104.

[29] M. Al-Ayyoub, S. B. Essa, and I. Alsmadi, "Lexicon-based sentiment analysis of Arabic tweets," *International Journal of Social Network Mining*, vol. 2, no. 2, pp. 101–114, Jan. 2015, https://doi.org/10.1504/IJSNM.2015.072280.

[30] M. Biltawi, G. Al-Naymat, and S. Tedmori, "Arabic Sentiment Classification: A Hybrid Approach," in *International Conference on New Trends in Computing Sciences*, Amman, Jordan, Oct. 2017, pp. 104–108, https://doi.org/10.1109/ICTCS.2017.24.

[31] R. T. Khasawneh, H. A. Wahsheh, I. M. Alsmadi, and M. N. AI-Kabi, "Arabic sentiment polarity identification using a hybrid approach," in *6th International Conference on Information and Communication Systems*, Amman, Jordan, Apr. 2015, pp. 148–153, https://doi.org/10.1109/IACS.2015.7103218.

[32] K. Elshakankery and M. F. Ahmed, "HILATSA: A hybrid Incremental learning approach for Arabic tweets sentiment analysis," *Egyptian Informatics Journal*, vol. 20, no. 3, pp. 163–171, Nov. 2019, https://doi.org/10.1016/j.eij.2019.03.002.

[33] S. Alzu'bi, O. Badarneh, B. Hawashin, M. Al-Ayyoub, N. Alhindawi, and Y. Jararweh, "Multi-Label Emotion Classification for Arabic Tweets," in *Sixth International Conference on Social Networks Analysis, Management and Security*, Granada, Spain, Oct. 2019, pp. 499–504, https://doi.org/10.1109/SNAMS.2019.8931715.

[34] O. Badarneh, M. Al-Ayyoub, N. Alhindawi, L. A. Tawalbeh, and Y. Jararweh, "Fine-Grained Emotion Analysis of Arabic Tweets: A Multi-target Multi-label Approach," in *12th International Conference on Semantic Computing*, Laguna Hills, CA, USA, Feb. 2018, pp. 340–345, https://doi.org/10.1109/ICSC.2018.00070.

[35] D. E. Salhi, A. Tari, and M.-T. Kechadi, "Sentiment Analysis Application on Twitter for E-reputation," in *6th International Conference on Image and Signal Processing and their Applications*, Mostaganem, Algeria, Nov. 2019, pp. 1–6, https://doi.org/10.1109/ISPA48434.2019.8966833.

[36] A. Baccouche, B. Garcia-Zapirain, and A. Elmaghraby, "Annotation Technique for Health-Related Tweets Sentiment Analysis," in *International Symposium on Signal Processing and Information Technology*, Louisville, KY, USA, Dec. 2018, pp. 382–387, https://doi.org/10.1109/ISSPIT.2018.8642685.

[37] R. Ismail, M. Omer, M. Tabir, N. Mahadi, and I. Amin, "Sentiment Analysis for Arabic Dialect Using Supervised Learning," in *International Conference on Computer, Control, Electrical, and Electronics Engineering*, Khartoum, Sudan, Dec. 2018, pp. 1–6, https://doi.org/10.1109/ICCCEEE.2018.8515862.

[38] M. E. M. Abo, N. A. K. Shah, V. Balakrishnan, and A. Abdelaziz, "Sentiment analysis algorithms: evaluation performance of the Arabic and English language," in *International Conference on Computer, Control, Electrical, and Electronics Engineering*, Khartoum, Sudan, Aug. 2018, pp. 1–5, https://doi.org/10.1109/ICCCEEE.2018.8515844.

[39] N. F. Alshammari and A. A. AlMansour, "Aspect-based Sentiment Analysis for Arabic Content in Social Media," in *International Conference on Electrical, Communication, and Computer Engineering*, Istanbul, Turkey, Jun. 2020, pp. 1–6, https://doi.org/10.1109/ICECCE49384.2020.9179327.

[40] N. F. Bin Hathlian and A. M. Hafezs, "Sentiment - subjective analysis framework for arabic social media posts," in *4th Saudi International Conference on Information Technology (Big Data Analysis) (KACSTIT)*, Riyadh, Saudi Arabia, Nov. 2016, pp. 1–6, https://doi.org/10.1109/KACSTIT.2016.7756073.

[41] M. Alassaf and A. M. Qamar, "Improving Sentiment Analysis of Arabic Tweets by One-way ANOVA," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 6, Part A, pp. 2849–2859, Jun. 2022, https://doi.org/10.1016/j.jksuci.2020.10.023.

[42] H. Al-Rubaiee, R. Qiu, and D. Li, "Identifying Mubasher software products through sentiment analysis of Arabic tweets," in *International Conference on Industrial Informatics and Computer Systems*, Sharjah, United Arab Emirates, Mar. 2016, pp. 1–6, https://doi.org/10.1109/ICCSII.2016.7462396.

[43] S. Tartir and I. Abdul-Nabi, "Semantic Sentiment Analysis in Arabic Social Media," *Journal of King Saud University - Computer and Information Sciences*, vol. 29, no. 2, pp. 229–233, Apr. 2017, https://doi.org/10.1016/j.jksuci.2016.11.011.

[44] S. Abuelenin, S. Elmougy, and E. Naguib, "Twitter Sentiment Analysis for Arabic Tweets," in *International Conference on Advanced Intelligent Systems and Informatics*, Cairo, Egypt, Sep. 2017, pp. 467–476, https://doi.org/10.1007/978-3-319-64861-3_44.

[45] K. M. Alomari, H. M. ElSherif, and K. Shaalan, "Arabic Tweets Sentimental Analysis Using Machine Learning," in *International Conference on Industrial, Engineering and Other Applications of*

Applied Intelligent Systems, Arras, France, Jun. 2017, pp. 602–610, https://doi.org/10.1007/978-3-319-60042-0_66.

[46] A. Mahmoud and T. Elghazaly, "Using Twitter to Monitor Political Sentiment for Arabic Slang," in *Intelligent Natural Language Processing: Trends and Applications*, K. Shaalan, A. E. Hassanien, and F. Tolba, Eds. New York, NY, USA: Springer, 2018, pp. 53–66.

[47] R. Bouchlaghem, A. Elkhelifi, and R. Faiz, "SVM based approach for opinion classification in Arabic written tweets," in *12th International Conference of Computer Systems and Applications*, Marrakech, Morocco, Nov. 2015, pp. 1–4, https://doi.org/10.1109/AICCSA.2015.7507153.

[48] L. S. Larkey, L. Ballesteros, and M. E. Connell, "Light Stemming for Arabic Information Retrieval," in *Arabic Computational Morphology: Knowledge-based and Empirical Methods*, A. Soudi, A. van den Bosch, and G. Neumann, Eds. Dordrecht, Netherlands: Springer, 2007, pp. 221–243.

[49] S. Khoja and R. Garside, *Stemming Arabic Text*. Bailrigg, England: Lancaster University, 1999.

[50] A. Al-Thubaity, Q. Alqahtani, and A. Aljandal, "Sentiment lexicon for sentiment analysis of Saudi dialect tweets," *Procedia Computer Science*, vol. 142, pp. 301–307, Jan. 2018, https://doi.org/10.1016/j.procs.2018.10.494.

[51] M. M. Altawaier and S. Tiun, "Comparison of Machine Learning Approaches on Arabic Twitter Sentiment Analysis," *International Journal on Advanced Science, Engineering and Information Technology*, vol. 6, pp. 1067–1073, Dec. 2016, https://doi.org/10.18517/ijaseit.6.6.1456.

[52] H. K. Aldayel and A. M. Azmi, "Arabic tweets sentiment analysis – a hybrid scheme," *Journal of Information Science*, vol. 42, no. 6, pp. 782–797, Dec. 2016, https://doi.org/10.1177/0165551515610513.

[53] A. Shoukry and A. Rafea, "A Hybrid Approach for Sentiment Classification of Egyptian Dialect Tweets," in *First International Conference on Arabic Computational Linguistics*, Cairo, Egypt, Apr. 2015, pp. 78–85, https://doi.org/10.1109/ACLing.2015.18.

[54] H. Abdellaoui and M. Zrigui, "Using Tweets and Emojis to Build TEAD: an Arabic Dataset for Sentiment Analysis," *Computacion y Sistemas*, vol. 22, no. 3, pp. 777–786, Sep. 2018, https://doi.org/10.13053/cys-22-3-3031.

[55] N. El-Naggar, Y. El-Sonbaty, and M. A. El-Nasr, "Sentiment analysis of modern standard Arabic and Egyptian dialectal Arabic tweets," in *Computing Conference*, London, UK, Jul. 2017, pp. 880–887, https://doi.org/10.1109/SAI.2017.8252198.

[56] R. Bouchlaghem, A. Elkhelifi, and R. Faiz, "Sentiment Analysis in Arabic Twitter Posts Using Supervised Methods with Combined Features," in *17th International Conference on Intelligent Text Processing and Computational Linguistics*, Konya, Turkey, Apr. 2016, pp. 320–334, https://doi.org/10.1007/978-3-319-75487-1_25.

[57] B. Brahimi, M. Touahria, and A. Tari, "Data and Text Mining Techniques for Classifying Arabic Tweet Polarity," *Journal of Digital Information Management*, vol. 14, no. 1, pp. 15–25, 2016.

[58] L. Al-Horaibi and M. B. Khan, "Sentiment analysis of Arabic tweets using text mining techniques," in *First International Workshop on Pattern Recognition*, Tokyo, Japan, Dec. 2016, vol. 10011, pp. 288–292, https://doi.org/10.1117/12.2242187.

[59] R. Bouchlaghem, A. Elkhelifi, and R. Faiz, "A Machine Learning Approach For Classifying Sentiments in Arabic tweets," in *6th International Conference on Web Intelligence, Mining and Semantics*, New York, NY, USA, Jun. 2016, pp. 1–6, https://doi.org/10.1145/2912845.2912874.

[60] Y. D. Setiyaningrum, A. F. Herdajanti, C. Supriyanto, and Muljono, "Classification of Twitter Contents using Chi-Square and K-Nearest Neighbour Algorithm," in *International Seminar on Application for Technology of Information and Communication*, Semarang, Indonesia, Sep. 2019, pp. 1–4, https://doi.org/10.1109/ISEMANTIC.2019.8884290.

[61] M. Alruily, "Towards Automatically Extracting Contextual Valence Shifters in Reviews of Saudi Universities," in *9th Symposium on Computer Applications & Industrial Electronics*, Malaysia, Apr. 2019, pp. 278–281, https://doi.org/10.1109/ISCAIE.2019.8743859.

[62] N. Al-Twairesh and H. Al-Negheimish, "Surface and Deep Features Ensemble for Sentiment Analysis of Arabic Tweets," *IEEE Access*, vol. 7, pp. 84122–84131, 2019, https://doi.org/10.1109/ACCESS.2019.2924314.

[63] A. Alotaibi and M. H. Abul Hasanat, "Racism Detection in Twitter Using Deep Learning and Text Mining Techniques for the Arabic Language," in *First International Conference of Smart Systems and Emerging Technologies*, Riyadh, Saudi Arabia, Nov. 2020, pp. 161–164, https://doi.org/10.1109/SMART-TECH49988.2020.00047.

[64] 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," *Social Network Analysis and Mining*, vol. 10, no. 1, Jul. 2020, Art. no. 53, https://doi.org/10.1007/s13278-020-00668-1.

[65] R. Baly *et al.*, "A Characterization Study of Arabic Twitter Data with a Benchmarking for State-of-the-Art Opinion Mining Models," in *Third Arabic Natural Language Processing Workshop*, Valencia, Spain, Apr. 2017, pp. 110–118, https://doi.org/10.18653/v1/W17-1314.

[66] G. Badaro, R. Baly, H. Hajj, N. Habash, and W. El-Hajj, "A Large Scale Arabic Sentiment Lexicon for Arabic Opinion Mining," in *EMNLP 2014 Workshop on Arabic Natural Langauge Processing*, Doha, Qatar, Oct. 2014, pp. 165–173.

[67] N. Al-Twairesh, H. Al-Khalifa, A. Al-Salman, and Y. Al-Ohali, "AraSenTi-Tweet: A Corpus for Arabic Sentiment Analysis of Saudi Tweets," *Procedia Computer Science*, vol. 117, pp. 63–72, Jan. 2017, https://doi.org/10.1016/j.procs.2017.10.094.

[68] S. M. Mohammad, M. Salameh, and S. Kiritchenko, "How Translation Alters Sentiment," *Journal of Artificial Intelligence Research*, vol. 55, pp. 95–130, Jan. 2016, https://doi.org/10.1613/jair.4787.

[69] R. M. Duwairi and I. Qarqaz, "Arabic Sentiment Analysis Using Supervised Classification," in *International Conference on Future Internet of Things and Cloud*, Barcelona, Spain, Aug. 2014, pp. 579–583, https://doi.org/10.1109/FiCloud.2014.100.

[70] R. M. Duwairi, "Sentiment analysis for dialectical Arabic," in *6th International Conference on Information and Communication Systems*, Amman, Jordan, Apr. 2015, pp. 166–170, https://doi.org/10.1109/IACS.2015.7103221.

[71] H. S. Ibrahim, S. M. Abdou, and M. Gheith, "MIKA: A tagged corpus for modern standard Arabic and colloquial sentiment analysis," in *2nd International Conference on Recent Trends in Information Systems*, Kolkata, India, Jul. 2015, pp. 353–358, https://doi.org/10.1109/ReTIS.2015.7232904.

[72] D. Mouheb, R. Albarghash, M. F. Mowakeh, Z. A. Aghbari, and I. Kamel, "Detection of Arabic Cyberbullying on Social Networks using Machine Learning," in *16th International Conference on Computer Systems and Applications*, Abu Dhabi, United Arab Emirates, Nov. 2019, pp. 1–5, https://doi.org/10.1109/AICCSA47632.2019.9035276.

[73] A. Al-Laith and M. Shahbaz, "Tracking sentiment towards news entities from Arabic news on social media," *Future Generation Computer Systems*, vol. 118, pp. 467–484, May 2021, https://doi.org/10.1016/j.future.2021.01.015.

[74] A. Al-Thubaity, M. Alharbi, S. Alqahtani, and A. Aljandal, "A Saudi Dialect Twitter Corpus for Sentiment and Emotion Analysis," in *21st Saudi Computer Society National Computer Conference*, Riyadh, Saudi Arabia, Apr. 2018, pp. 1–6, https://doi.org/10.1109/NCG.2018.8592998.

[75] F. Sadat, F. Mallek, M. Boudabous, R. Sellami, and A. Farzindar, "Collaboratively Constructed Linguistic Resources for Language Variants and their Exploitation in NLP Application – the case of Tunisian Arabic and the Social Media," in *Workshop on Lexical and Grammatical Resources for Language Processing*, Dublin, Ireland, Aug. 2014, pp. 102–110, https://doi.org/10.3115/v1/W14-5813.

[76] N. El-Makky *et al.*, *Sentiment Analysis of Colloquial Arabic Tweets*. Alexandria, Egypt: Alexandria University, 2014.

[77] M. F. Alhamid, S. Alsahli, M. Rawashdeh, and M. Alrashoud, "Detection and Visualization of Arabic Emotions on Social Emotion Map," in *International Symposium on Multimedia*, Taichung, Taiwan, Dec. 2017, pp. 378–381, https://doi.org/10.1109/ISM.2017.76.

[78] W. M. S. Yafooz, E. A. Hizam, and W. A. Alromema, "Arabic Sentiment Analysis on Chewing Khat Leaves using Machine Learning and Ensemble Methods," *Engineering, Technology & Applied Science Research*, vol. 11, no. 2, pp. 6845–6848, Apr. 2021, https://doi.org/10.48084/etasr.4026.

[79] M. Mahyoob, J. Algaraady, M. Alrahiali, and A. Alblwi, "Sentiment Analysis of Public Tweets Towards the Emergence of SARS-CoV-2 Omicron Variant: A Social Media Analytics Framework," *Engineering, Technology & Applied Science Research*, vol. 12, no. 3, pp. 8525–8531, Jun. 2022, https://doi.org/10.48084/etasr.4865.

[80] I. A. Kandhro, S. Z. Jumani, F. Ali, Z. U. Shaikh, M. A. Arain, and A. A. Shaikh, "Performance Analysis of Hyperparameters on a Sentiment Analysis Model," *Engineering, Technology & Applied Science Research*, vol. 10, no. 4, pp. 6016–6020, Aug. 2020, https://doi.org/10.48084/etasr.3549.