Opinion mining for Arabic dialect in social media data fusion platforms: A systematic review

Hani D. Hejazi, Ahmed A. Khamees

Faculty of Engineering &IT, The British University in Dubai, UAE

Emails: hani.hejazi@gmail.com; khamisos@gmail.com

Abstract

The huge text generated on social media in Arabic, especially the Arabic dialect becomes more attractive for Natural Language Processing (NLP) to extract useful and structured information that benefits many domains. The more challenging point is that this content is mostly written in an Arabic dialect with a big data fusion challenge, and the problem with these dialects it has no written rules like Modern Standard Arabic (MSA) or traditional Arabic, and it is changing slowly but unexpectedly. One of the ways to benefit from this huge data fusion is opinion mining, so we introduce this systematic review for opinion mining from Arabic text dialect for the years from 2016 until 2019. We have found that Saudi, Egyptian, Algerian, and Jordanian are the most studied dialects even if it is still under development and need a bit more effort, nevertheless, dialects like Mauritanian, Yemeni, Libyan, and somalin have not been studied in this period. Many data fusion models that show a good result is the last four years have been discussed.

Keywords: Data Fusion; Arabic dialect; Natural language processing; Opinion mining; Systematic Review.

1. Introduction

1.1. Background

The evolution in using social media applications like Twitter, Instagram, and Facebook has developed into a huge number of users who contribute to building a large volume of opinions and reviews about different topics [1]–[12]. Humans tend to share their thoughts, feelings, and opinions according to their experiences, observation, and understanding of events, services, or products [13]–[21]. Their impressions or point of view may be expressed to be positive opinions, negative or neutral. These opinions can help decision-makers, marketers, business owners, or service providers to identify their customer's interests in a particular product, service, or event. The vast amount of posted text on social media raises the need to have an analysis tool to extract the opinions and classify them. This process of investigation through the user's emotions and views is called Opinion Mining [22]. The Arabic language is one of the six official languages used in the United Nations and the official language of 22 countries [23], [24]. Arabic is the speaking language of more than 350 million people all over the world. In 1948, at the 3rd General Conference of UNESCO held in Lebanon, Arabic was declared to become the third language to work with, besides English and French, in the governing body meetings in the Arabic countries.

Since then, the Arabic language gain more attention when the International Day of the Arabic language is announced to be celebrated on December 18 of each year, the day when the General Assembly of the United Nations adopted Arabic as the sixth official language of the Organization in 1973 [25]. In the practical world of the internet, Arabic is considered a complex and rich language. With more than 226 million Arabic users on the internet and over 141 million users on Facebook. Moreover, Arabic is the 4th most language used on the web [1], [23], [26]. There are 28 Arabic alphabet letters with different shapes of each one of them according to the letter placed in the written word. Writing Arabic is implemented from right to left [23].

Doi : https://doi.org/10.54216/FPA.090101
Received: April 10, 2022 Accepted: September 22, 2022

There are 22 Arabic Dialects (AD) that represent the official Arabic-speaking countries [27], but the Modern Standard Arabic (MSA) is the formal language that is used to express formal communication such as in the news, books magazines, and official statements. However, the Holy Quran language is considered Classical Arabic [28]. MSA and DA could be written in either Arabic letters or using Romanian ones (Arabizi). Due to the complexity of the Arabic language, Arabizi is used to outline the Arabic words using a combination of Latin characters, numbers, and signs [29]. MSA and DA can be classified into six groups: Egyptian (used in Egypt), Khaliji (used in the Arabic Gulf area), Shami (Levantine) spoken in (Jordan, Lebanon, Palestine, and Syria), Maghrebi (northern Africa), Sudanese Arabic and Iraqi [30]. Each one of these groups includes some number of sub-Arabic vernaculars [28]. Arabic is a rich language in morphological and vocabulary. It can be exposed in many varieties. The formal-informal dialects of MSA and DA usage become more critical for communication after the emergence of social networks and the increasing number of users that result in a huge volume of the unstructured format of the text published on the internet. MSA is a type of Arabic language form that is not considered a native one of any country, and it is much different from the informal DA. Moreover, MSA and DA complexity make it difficult to identify what those sections of text do represent. On the other hand, the process of Arabic Opinion Mining has several challenges due to the limited number of Natural Language Processing (NLP) tools and the different nature of contents found on the web, as most of it is written in a dialectal scheme1 [31]. NLP is defined as a collection of approaches that a computer employs to be capable of analyzing or understanding the unprocessed natural spoken language between people through extracting grammar and meaning from the input [32]. Several numbers of studies have been done to examine the English language using different NLP tools to extract OM from the text published on social media. Whereas, to the Arabic language, less attention is made due to its complexity. In this paper, we are trying to highlight the most recent studies accomplished and published over the last 4 years (2016 to 2019). Our research is trying to investigate the following questions:

1) What is the most Arabic dialect been researched recently in opinion mining?
2) What is the algorithm used to evaluate opinion mining of the Arabic dialect?
3) What are the most publishing journals and conferences in opinion mining of the Arabic dialect?
4) What are the gaps and shortages in Arabic dialect research in opinion mining?

In this paper, a systematic review has been conducted on 223 articles. In the first part, we gave a brief about the Arabic Language and its different text classifications. Then, we addressed the challenges when processing the Arabic language. After that, we conducted a literature review that would answer our research questions mentioned above. Next, we explained our methodology and approach to filtering out the intended papers under investigation — finally, we have shown the findings of the study.

1.2. Challenges in Arabic language processing

The process of analyzing Arabic language MSA and DA presents several challenges, which can be briefed in the following points:

MSA vs DA: MSA differs from the more commonly local DA used by people either in their daily interactions or over the internet. It's more likely seen that the main educated people who speak the Arabic language can understand MSA, but this is not fundamentally correct regarding the other DA. DA varies from MSA phonologically, morphologically, syntactically, and orthographically up to certain limitations as well. Indeed, building so many lexicons for those several dialects is quite challenging since data fusion transcription is time-consuming and expensive. It can be said that each DA or set of DA can be dealt with as a separate type of language. By contrast, MSA has a lot of resources [33].

- Lack of datasets: Most of the available datasets open for public use are small in measure of size. Having large sets of Arabic datasets is crucial for developers to build reliable OM tools to analyze the Arabic text. Moreover, compared to the English language, fewer MSA and DA lexicons are available in Arabic. The lack of such datasets hinders the research on similar topics using NLP tools such as dependency parser and Part of Speech POS tagger [31].
- Names can mislead: Some Arabic names of people or figures can be treated as adjectives, this can be confusing for OM.
- Negation ambiguity some words can hold the opposite meaning of what it does mean. These types of words can be used to get a funny conversation that was maybe irritating to others. Avoiding such

Doi : https://doi.org/10.54216/FPA.090101
Received: April 10, 2022 Accepted: September 22, 2022

ambiguity can produce a robust OM tool. Building a list of negation words can usually handle this key concern.

- Arabic Morphologically: Arabic is a rich morphological language where the syntax is expressed at the word base level. This Arabic base of a word can drive thousands of other surface words. A word in English can be shown in many words in Arabic [31].
- Using Arabizi: Using informal Arabic on internet content can lead to some other challenges. Using Latin letters within the Arabic word may confuse OM analysis tools since there is no standard to deal with these words that varies depending on the way users write them [31].
- Processing Performance: Many NLP tools have been developed to process MSA. However, these tools have limited performance over DA. (e.g., state of an art morphological analyzer of MSA has only 60% of coverage over Arabic Levantine dialect forms of verbs [33].
- Popularity: The rate of adding new words to the different categories of DA is more than it is for MSA since the DA is widely used and more popular than it is in MSA [33]. So, the challenges regarding Arabic content published on social media can be summarized as follows [34].
- Unstructured text.
- Orthographic typing mistakes.
- Using slang language.
- Many colloquial abbreviations and expressions are used on Twitter due to the limited number of characters.
- Inconsistencies in Spelling.
- No Arabic capital letters.
- Using emoticons.
- The tendency of repeating some letters when writing to express feelings. DA requires special handling compared to MSA due to DA several variations in the structure of the word.

2. Literature review

2.1. Opinion Mining (OM)

Opinion Mining (OM), also called Sentiment Analysis (SA), is a field of study that focuses on analyzing people's attitudes, emotions, and evaluations towards different entities such as services, products, individuals, organizations, topics, and events, issues along with their attributes. OM is considered very important yet has a large space of problems to determine if a word, sentence, or document expresses negative or positive feelings [35]. OM is the task of identifying emotions behind the attitude or behavior of a person through a computational process using his or her text as input. Then we can be able to identify whether a person's (e.g. writer) attitude towards a specific topic is positive, negative or neutral [22]. Recently, the number of published texts on social media is extensively increased. For instance, in 2017, many comments posted on Facebook every minute were over 500 thousand comments, and around 350 thousand tweets were sent using Twitter. Hence, using social media published content as an input to NLP tools shall reflect the opinion of the public people [36]. OM over social media in recent studies focused on different topics (i.e., Politics, Marketing), other OM focused on analyzing opinions on products or services reviews (i.e., Pizza industry, Phone companies). Other OM tools are used to monitor stock market performance. Gathering hate speech has also been studied to detect any hate posts based on Gender, Color Religious views or Racist kind of speech that are offending individuals of particular tribe [37]. But most of these studies investigated English language and less work been addressing DA or even MSA.

OM fields that develop emotions like anger, sadness, or happiness, and identifying these moods in the audio, video, or written texts are important fields in artificial intelligence tools and NLP. The Arabic language varies in its linguistic methods and art can be expressed in many types, such as poetry, prose, criticism, etc. these arts on MSA and DA involve several OM structures that gave deeper look at the meaning of the words [38]. Arabic Language OM includes tasks that include morphology phonetics, POS tagging, opinion mining, subjective analysis, Named Entity Recognition (NER), and annotation opinion manually using corpus or lexicons [39]. OM is performed using either one of the following two techniques: Rule-Based (RB) or Machine Learning (ML) classifiers, where ML algorithms are used to observe emotions and detect opinions [39]. Many datasets were used to evaluate several proven classifiers used in OM, such as Support Vector Machine (SVM), Naive Bayes (NB), K-Nearest Neighbor (KNN), and others [28].

2.2. Opinion approaches

Doi : https://doi.org/10.54216/FPA.090101
Received: April 10, 2022 Accepted: September 22, 2022

Existing OM approaches can be divided into three main approaches, Lexicon Based approach, the ML-based approach, or a hybrid approach that combines both Lexicon-based and ML approaches [40].

2.2.1 Lexicon-based methods

OM using lexicons to detect attitudes, feelings, and emotions includes a wide range of phrases and words that address opinion whether it holds a positive, negative, or neutral meaning. This kind of lexicon helps in identifying the actual meanings of Arabic linguistics parts of a text [39]. According to [40] Symbolic orientation is used to highlight opinions within the text using opinion lexicons; these lexicons are a list of phrases or words combined with a positive or negative opinion. Part of these lexicons gives a score to indicate how strong its class is. This technique shall provide an overall opinion as a sum of these scores. OM using lexicons can be manually created or automatically.

Lexicon-based approaches can be constructed for general-purpose lexicons or domain-specific ones. According to [40] Found the lexicon-based approach may result in a low percentage of recall in entity-level OM. This approach needs a huge amount of training data. However, there are important advantages to using Lexicon-based approaches, once the lexicons are built, then there would be no need to train the data.

2.2.2 Machine learning (ML) based methods

The ML-based approach can be classified into two sub-approaches: unsupervised and supervised, the accuracy of which is measured by how well they agree with human judgments [41]–[48]. The unsupervised technique is calculated depending on the polarity terms of the lexicon [11], [49]–[55]. Such lexicons may have some initial seeds, but several numbers of these lexicons are not domain-specific, which in turn are not useful for DA language. Extending such lexicons could be corpus-based or dictionary-based. Unlike the unsupervised technique, a supervised method mainly depends on ML. To classify the polarity of input data, some sets of data are collected for those labeled ones, and some other sets of features are extracted for training the ML classifier. The extracted features can be any handcrafted type of features like POS tagging, emoticons, lemmatized or stemmed. Then both extracted labels and features are input to an ML algorithm that is important in building a model for the ML classifier. Finally, labels are assigned to the ML classifier model over the text input extracted features. It has been investigated by [36] that the accuracy of the ML approach can outperform the lexicon-based approach by using a well-trained corpus and selecting a good set of features.

However, in [36] research on DA language, mistakes could not be handled like spelling and concatenation words by lemmatization. [40] defined ML as a computer able to learn itself the way to make decisions depending on previous experiences and available data (Training data). The decisions made by ML could be a prediction or classification for the new data, this new data is labeled (i.e. by experts) so that the learning algorithm uses them in classification. This learning algorithm task is called supervised classification. ML mainly counts on the training data used to train ML algorithms with many features, like stylistic, frequency co-occurrence, and more. These collected features were then used to test algorithm accuracy using a sample of the annotated data. This approach avoids any ambiguity in Arabic because it is a language-independent approach. [56], [57] studied test the data judged (labeled) by human experts to check how much it affects the accuracy of trained ML classifiers. They found that whenever they increase the training data, the accuracy shall, in turn, be improved. ML approach includes several statistical methods to build lexicons datasets, like NB, SVM, KNN, and decision trees. [40] found that these methods tend to have a high level of accuracy in detecting the OM of the input textual data. But the major drawback of these ML methods is related to the domain, trained data should be domain-specific so it does not work as well as it is supposed to on different text other than those trained [40]. A different number of supervised and unsupervised ML classifiers are used to evaluate the accuracy of Arabic datasets. Namely, we have investigated the following:

2.2.2.1 Naive Bayes (NB)

According to [50], NB is defined as a supervised classifier that focuses on training datasets with label tests. Then, the unlabeled dataset is tested using this classifier to determine OM. [58] defined NB as a probabilistic-based classifier. NB classifier algorithm can be applied to classify text polarity (OM) independently with strong rules that can be accurate, fast, simple, and effective. Thus, the NB algorithm can review problems that associate objects of the discrete type. To perform NB method text classification, features need to be extracted from the input text. Then features are saved in a vector for analysis, this is called (Term Vector TV or Feature Vector FV). This TV generation depends on the unique words from the

Doi : https://doi.org/10.54216/FPA.090101
Received: April 10, 2022 Accepted: September 22, 2022

trained dataset. As [58] mentioned, NB can be classified into two categories, the multinomial and multivariate models. In [59], researchers conducted OM on the Saudi dialect from Twitter, and a large set of text (corpus) was annotated manually to exclude any incorrect ones. This corpus was validated using NB and SVM classifier algorithms. Results showed that the NB algorithm outperformed other demonstrated methods and gained up to 91% accuracy revealing the importance of embedding words technique to build the lexicon. [60] implemented an NB classifier in MSA resources on Gulf dialects. The authors used the NB classifier due to its flexibility to add new features to the corpus and change the scores of OM. Results indicated that using MSA over Gulf dialects resources is useless because of some structural differences that made some conflicts between these two types of dialects.

2.2.2.2 Support Vector Machine (SVM)

[37] Defined SVM as a binary type of classifier that data samples are assumed to be clearly distinct from each other. SVM tries to look for the best hyper-plan that increases class margin. And divides the textual data to detect OM as positive or negative. This type of classifier is widely used in cyberbullying. [61] investigated four types of supervised approaches to detect OM over a dataset of reviews manually collected from the Jeeran website written in MSA. Results showed that the SVM algorithm got the highest score of classification accuracy equal to 92.3%. [62] studied the Jordanian dialect through gathering posts from formal Facebook pages belonging to the telecommunication companies in Jordan like Zain, Orange, and Umniah.

Data have been manually collected and then classified using four types of supervised classifiers. The algorithm of the SVM classifier performed the best compared to the other ones in terms of F-measure and accuracy. In [63] researchers collected Facebook comments written in MSA and Moroccan DA to analyze the OM about Moroccan elections that took place in 2016. The dataset was evaluated using several classifiers, where SVM obtained the best-developed algorithm model regardless of the configuration of the weighting or extraction.

2.2.2.3 Based approach

This type of method depends primarily on using polarity over lexicons. Users should tag words with positive, negative, or neutral labels. [39] Deriving Arabic words into their original form would help to set many clear rules that carry the opinion mining. Thus, [39] can assign scores based on these rules to capture the level of positive opinions or negative ones within the sentences. To enhance the accuracy of RB classification [39] used a mathematical Rough set theory (RST) tool to reduce the high extent of the vector feature.

2.2.2.4 Long short-term memory (LSTM)

LSTM is a deep learning approach used to detect OM objects. It focuses on finding the relationship between sequential input elements. LSTM is another kind of classifier of RNN. This classifier aims to process data on words found on Twitter to detect short and long-term relations between the current and previous words. Then, LSTM combines them in one output representing them in one output. Finally, the algorithm output is used as a decision layer to detect OM. LSTM is used to overcome problems in RNN using memory cells, preventing the forget status when multiple iterations are done [32]. [64] performed OM on Twitter posts written in the Saudi dialect and gathered in the SDCT dataset. The proposed model was evaluated by using LSTM and SVM models. Results showed that LSTM has better performance than the SVM algorithm. [32] Proposed a framework that improves OM. The research was demonstrated on a Moroccan corpus called MSAC gathered from the Arabic reviews and comments on Facebook page and Twitter. Results show that RNN outperforms other classical approaches like SVM and NB.

2.2.2.5 K-Nearest Neighbor Algorithm (KNN)

K-NN is a type of ML algorithm where functions are locally approximated, and computation is delayed until classification is accomplished. K-NN is a method used to do regression and classification where the input has k number of examples trained closest to the feature's space. The output of K-NN depends on whether classification or regression is used. K-NN is a useful approach to give neighbors a degree of the weight of contribution depending on how much nearer or distinct it is (nearer gains more weight). [Wikipedia https://tinyurl.com/c3avfz6] [65] Observed OM on data from a telecommunications company Twitter written

Doi : https://doi.org/10.54216/FPA.090101
Received: April 10, 2022 Accepted: September 22, 2022

in Sudanese dialect. Researchers conducted several ML approaches to compare the performances of each of them. Results found that using KNN gave the best performance in terms of accuracy with (k=2), which equals 92%. [66] explored the different effects of hierarchal classifiers based on their performances to other models. Researchers found that KNN hierarchical classification HC got the best accuracy, it obtained 74.64%.

2.2.2.6 Other

Other methods were found too like Polynomial Networks (PNs), acoustic model, logistic, regression, 4.1.3 Decision Tree (DT), CBOW, SenticNet, Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN).

2.2.2.7 Hybrid Approach

[39] Found that using a combination of two ML techniques (supervised and unsupervised) like SVM and KNN where SVM works better for large sets of data, and KNN is better when targeting the small size of the corpus. However, when using the unsupervised approach, there will be no human action for labeling data; therefore, function learning counts on searching for any patterns that may be found within the unlabeled data. Some articles used a hybrid approach where many methods were applied together, and the results were selected upon a certain algorithm, like using rule-based and SVM and then building a model to choose between the results according to specific criteria [67]–[69].

3. METHODOLOGY

3.1. Data sources and search strategies

We have chosen Emerald, Google Scholar, IEEE, ProQuest, ScienceDirect, Springer, Taylor, and Wiley databases as a source for the review papers since it includes most of the recently published papers and is well-indexed. Since we are focusing on opinions from Arabic dialects written in social media in the last four years. The search criteria were made to include “opinion mining” as a mandatory phrase also “social media” as mandatory too while it was a choice between “Arabic dialect" and ("Arabic dialects" since the result was changed when we use each one alone and papers year starting from 2016 till 2019, the exact search criteria as described in Table 1.

Table 1: Keyword search

Keyword search
"Opinion Mining" & "Arabic dialects" & “social media”
"Opinion Mining " & "Arabic dialects" & “Tweets”
"Opinion Mining " & "Arabic dialects" & “Facebook”
"Opinion Mining " & "Arabic dialect" & “social media”
"Opinion Mining " & "Arabic dialect" & “Tweets”
"Opinion Mining " & "Arabic dialects" & “Facebook”

A total of 223 articles were returned from different journals, conferences, Ph.D. thesis, books, or patents. Each with separate citation count, so the next step was to evaluate the papers to include the related ones. After reviewing the papers we have excluded the papers that are review papers, papers that experiment with MSA only, and Ph.D. Thesis, books, or unavailable papers due to their cost and membership requirements. After that in each paper, we go deep to get the best approach that is used and shows the best opinion mining results since most papers tried different techniques for opinion mining and compare them to find the best one we have found the following techniques and classifiers: SVM, Naive Bayes, LSTM, CNN, KNN, RNN, Polynomial, Networks (PNs), rule-based, lexicon-based, acoustic model, logistic regression, CBOW,

Doi : https://doi.org/10.54216/FPA.090101
Received: April 10, 2022 Accepted: September 22, 2022

decision tree classifier, hybrid, SenticNet, and Maximum Entropy. Also, we state the dataset used and its source; we divided the dataset into four groups:

- **Tweets:** for datasets collected from twitter in certain conditions like location or accounts, since tweet size is limited we think that certain features will be common between the tweets where users have limited words to express their idea and sometime user needs to rephrase their sentence to fit in one tweet.
- **Facebook comments and posts:** for datasets extracted from Facebook comments or long posts, since this has fewer word limits we think it will be easier for a user to express his ideas without rephrasing.

- **User reviews:** for datasets extracted from books, restaurants, movies, trips, or location reviews, we think these reviews will have some common way of expressing ideas since it is centered on certain features of the reviewed subjects.
- **Others:** for datasets that are from the ready corpus or previous datasets, or if the dataset source is not mentioned clearly. Also, we have extracted the Arabic dialects that have been studied in each paper; in some papers that discussed more than one dialect we have found the following dialects: Jordanian, General (many dialects), Egyptian, Saudi, Gulf, Levantine, MSA, Arabizi, Syrian, Moroccan, Algerian, Sudanese, Iraqi, Lebanese, Tunisian, UAE, and North Africa.

3.2. Inclusion/exclusion criteria

A total of 288 articles included these keywords. Among these, it was found that 65 articles were repeated, and so they were eliminated. A total of 223 papers were now remaining for review. The way these articles are distributed concerning the databases they are a part of is shown in Table 2. The researcher checked the inclusion and exclusion criteria for each study [70]–[75]. One hundred three research articles all met the inclusion criteria; hence, the analysis was based on these studies. Figure 1 depicts the systematic review process and the total articles found at each stage.

Table 2: The total number of articles after removing the duplicates.

Database Frequency	Frequency
Emerald	16
Google Scholar	98
IEEE	25
ProQuest	15
ScienceDirect	21
Springer	29
Taylor	6
Wiley	13
Total	223

After reviewing the articles we have eliminated the unrelated articles that don't conduct the opinion mining specifically or not related to any Arabic dialect, some articles study only MSA, it does only a review of other articles and work, or don't conduct a specific method for evaluation, so after this process, we get 103 related articles out of total 223 articles, 120 articles were not related as described in Table 3. The articles selected in this review study for a critical review should be used to complete the inclusion and exclusion criteria presented in Table 4.

Table 3: Reviewed articles count.

Articles	Articles count
All Articles	223
Related	103
Unrelated	120

Doi : https://doi.org/10.54216/FPA.090101
Received: April 10, 2022 Accepted: September 22, 2022

Table 4; Inclusion and exclusion criteria.

Inclusion Criteria	Exclusion Criteria
Must involve opinion mining.	Articles without opinion mining.
Must involve the Arabic dialect.	Articles with opinion mining but without Arabic dialect.
Must be written in the English language.	Articles published in languages other than English.
Must be published between 2016 and 2019.	

3.4. Quality assessment

Apart from the inclusion and exclusion criteria, quality assessment of the selected articles is another factor that may be included [74], [76]–[82]. A quality assessment checklist that had 9 criteria was used to evaluate the quality of the articles that would subsequently be examined (N=103). The quality assessment checklist is shown in Table 5. The checklist was not developed to function as a means of disapproving the studies performed by any author [83]. The criteria presented by [83] served as the foundation on which the checklist was developed. Scores were awarded to each question using the three-point scale, where 1 point was awarded to "Yes", 0 points were awarded to "No," and 0.5 points were awarded to "Partially". Hence, a score of 0 to 9 could be attained by each study, where the high degree to which the research questions are answered by the study is represented by a high score. Table 6 demonstrates the quality assessment findings for the 103 studies. It is demonstrated in the findings that all studies have met the quality assessment criteria, thus showing that each of these studies is qualified for being employed in further analysis.

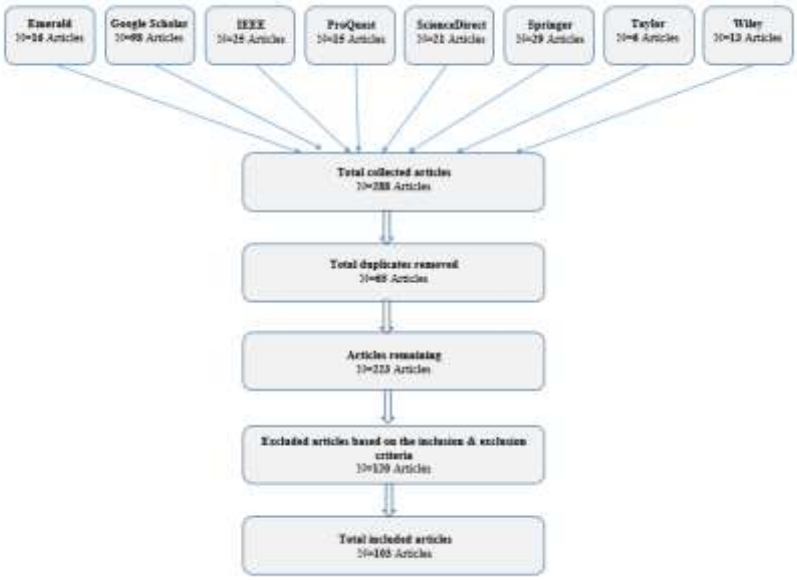

Figure 1: Systematic review process.

Table 5: Quality assessment checklist.

#	Question
1	Are the research aims clearly specified?
2	Was the study designed to achieve these aims?
3	Are the variables considered by the study clearly specified?
4	Is the study context/discipline clearly specified?
5	Are the data collection methods adequately detailed?
6	Does the study explain the reliability/validity of the measures?
7	Are the statistical techniques used to analyze the data adequately described?

Doi : https://doi.org/10.54216/FPA.090101
Received: April 10, 2022 Accepted: September 22, 2022

8	Do the results add to the literature?
9	Does the study add to your knowledge or understanding?

Table 6: Quality assessment results.

Study	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Total	Percentage
A1	1	1	1	1	1	1	1	1	1	9	100%
A2	1	1	1	1	1	1	1	1	1	9	100%
A3	1	1	1	1	1	1	1	1	1	9	100%
A4	1	1	1	1	1	1	1	1	1	9	100%
A5	1	1	1	1	1	1	1	1	1	9	100%
A6	1	1	1	1	1	1	1	1	1	9	100%
A7	1	1	1	1	1	1	1	1	1	9	100%
A8	1	1	1	1	1	1	1	1	1	9	100%
A9	1	1	1	1	1	1	1	1	1	9	100%
A10	1	1	1	1	1	1	1	1	1	9	100%
A11	1	1	1	1	1	1	1	1	1	9	100%
A12	1	1	1	1	1	1	1	1	1	9	100%
A13	1	1	1	1	1	1	1	1	1	9	100%
A14	1	1	1	1	1	1	1	1	1	9	100%
A15	0.5	1	1	1	1	0.5	1	1	1	8	89%
A16	1	1	1	1	1	1	1	1	1	9	100%
A17	1	1	1	1	1	1	1	1	1	9	100%
A18	1	1	1	1	1	1	1	1	1	9	100%
A19	1	1	1	1	1	1	1	1	1	9	100%
A20	1	1	1	1	1	1	1	1	1	9	100%
A21	1	1	1	1	1	1	1	1	1	9	100%
A22	1	1	1	1	1	1	1	1	1	9	100%
A23	1	1	1	1	1	1	1	1	1	9	100%
A24	1	1	1	1	1	1	1	1	1	9	100%
A25	1	1	1	1	1	1	1	1	1	9	100%
A26	1	1	1	1	1	1	1	1	1	9	100%
A27	1	1	1	1	1	1	1	1	1	9	100%
A28	1	1	1	1	1	1	1	1	1	9	100%
A29	1	1	1	1	1	1	1	1	1	9	100%
A30	1	1	1	1	1	1	1	1	1	9	100%
A31	1	1	1	1	1	1	1	1	1	9	100%
A32	1	1	1	1	1	1	1	1	1	9	100%
A33	1	1	1	1	1	1	1	1	1	9	100%
A34	1	1	1	1	1	1	1	1	1	9	100%
A35	1	1	1	1	1	1	1	1	1	9	100%
A36	1	1	1	1	1	1	1	1	1	9	100%
A37	1	1	1	1	1	1	1	1	1	9	100%
A38	1	1	1	1	1	1	1	1	1	9	100%
A39	1	1	1	1	1	1	1	1	1	9	100%
A40	1	1	1	1	1	0	1	1	1	8	89%
A41	1	1	1	1	1	1	1	1	1	9	100%
A42	1	1	1	0	1	1	1	1	1	8	89%
A43	1	1	1	0	1	1	1	1	1	8	89%
A44	1	1	1	0	1	1	1	1	1	8	89%
A45	1	1	1	1	1	1	1	0.5	0.5	8	89%
A46	1	1	1	0	1	1	1	1	1	8	89%
A47	1	1	1	0	1	1	1	1	1	8	89%
A48	1	1	1	0.5	1	1	1	1	0.5	8	89%
A49	1	1	1	0	1	1	1	1	1	8	89%
A50	1	1	1	1	1	1	1	1	1	9	100%

Doi : https://doi.org/10.54216/FPA.090101
Received: April 10, 2022 Accepted: September 22, 2022

A51	1	1	1	0.5	0.5	0.5	1	1	0.5	7	78%
A52	1	1	1	0	1	1	1	1	1	8	89%
A53	1	0	1	1	0	1	0.5	1	0.5	6	67%
A54	1	0.5	1	1	0.5	1	1	1	1	8	89%
A55	1	1	1	1	1	1	1	1	1	9	100%
A56	1	1	1	1	1	1	1	1	1	9	100%
A57	1	1	1	1	1	1	1	1	1	9	100%
A58	1	0.5	1	0	1	1	0	1	0.5	6	67%
A59	1	1	1	1	1	1	1	1	1	9	100%
A60	1	1	1	1	1	1	1	1	1	9	100%
A61	1	1	1	1	1	1	1	1	1	9	100%
A62	1	1	1	1	1	1	1	1	1	9	100%
A63	1	0	1	1	1	1	1	1	1	8	89%
A64	1	1	1	0	1	1	1	1	1	8	89%
A65	0.5	1	1	0.5	0.5	0.5	1	1	0.5	6.5	72%
A66	1	1	1	1	1	1	1	1	1	9	100%
A67	1	1	1	1	1	1	1	1	1	9	100%
A68	1	1	1	0	1	1	1	1	1	8	89%
A69	1	1	1	1	1	1	1	1	1	9	100%
A70	1	1	1	1	1	1	1	1	1	9	100%
A71	1	1	1	1	1	1	1	1	1	9	100%
A72	1	1	1	1	1	1	1	1	1	9	100%
A73	1	1	1	1	1	1	1	1	1	9	100%
A74	1	1	1	0	0.5	1	0.5	1	1	7	78%
A75	1	1	1	1	1	1	1	1	1	9	100%
A76	1	1	1	0	1	1	1	1	1	8	89%
A77	1	1	1	1	1	1	1	1	1	9	100%
A78	1	1	1	1	1	1	1	1	1	9	100%
A79	1	1	1	1	1	1	1	1	1	9	100%
A80	1	1	1	1	1	1	1	1	1	9	100%
A81	1	1	1	1	1	1	1	1	1	9	100%
A82	1	1	1	1	1	1	1	1	1	9	100%
A83	1	1	1	1	1	1	1	1	1	9	100%
A84	1	1	1	1	1	1	1	1	1	9	100%
A85	1	1	1	1	1	1	1	1	1	9	100%
A86	1	1	1	1	1	1	1	1	1	9	100%
A87	1	1	1	1	1	1	1	1	1	9	100%
A88	1	0.5	1	0	0.5	1	1	0.5	1	6.5	72%
A89	1	1	1	1	1	1	1	1	1	9	100%
A90	1	1	1	0.5	1	0.5	1	1	1	8	89%
A91	1	1	1	1	1	1	1	1	1	9	100%
A92	1	1	1	1	1	1	1	0	1	8	89%
A93	1	1	1	1	1	1	1	1	1	9	100%
A94	1	1	1	1	1	1	1	1	1	9	100%
A95	1	1	1	1	1	1	1	1	1	9	100%
A96	1	1	1	0	1	1	1	1	1	8	89%
A97	1	1	1	1	1	1	1	1	1	9	100%
A98	1	1	1	1	0	1	1	1	1	8	89%
A99	1	1	1	1	1	1	1	1	1	9	100%
A100	1	1	1	1	1	1	1	1	1	9	100%
A101	1	1	1	1	1	1	1	1	1	9	100%
A102	1	1	1	1	1	1	1	1	1	9	100%
A103	1	1	1	1	1	1	1	1	1	9	100%

Doi : https://doi.org/10.54216/FPA.090101
Received: April 10, 2022 Accepted: September 22, 2022

4. Results and Discussions

4.1. Methods used

We have covered the methods used in each paper and the methods that demonstrated the best opinion mining results for the Arabic dialect on the specified dataset, so we found that these methods appear to give the best results in the related papers: SVM, Naive Bayes, LSTM, CNN, KNN, RNN, Polynomial Networks (PNs), Rule-Based, Lexicon-Based, Acoustic Model, Logistic Regression, CBOW, Decision Tree classifier, Hybrid, SenticNet, Maximum Entropy. The result is illustrated in Table 7, and it showed that SVM provided the best results in 31 articles and then lexicon-based in 24 articles, all results are listed in Table 7.

Table 7. The total number of articles that demonstrated OM of Arabic language in the listed OM methods.

Method	Articles count
SVM	31
Lexicon-based	24
Hybrid	12
Naive Bayes	7
LSTM	4
CNN	4
KNN	4
Decision Tree	3
Rule-based	2
Logistic regression	2
SenticNet	2
RNN	1
Polynomial Networks (PNs)	1
Acoustic model	1
CBOW	1
Maximum Entropy	1

Some Papers compare more than five methods to extract the opinion; we include the one that produced the best results so that we can cover the best methods used and tested for opinion mining in the Arabic dialect.

4.2. Arabic Dialects

Many dialects were studied in the related papers, we have collected them to analyze the top studded dialect and least ones so we can show the dialects that need more elaboration and the top covered ones so other researchers can build on top of the other articles. The results show that the most studded dialect was Egyptian with 13 articles, Saudi with 12 articles, Algerian with 9 articles, then Jordanian with 8 articles. Tables 8 illustrates the dialect studied and the number of papers that discuss these dialects were a general is included for the papers that collect general Arabic dialect without biasing and not mentioning the main dialect in the article, some papers discuss the MSA with the dialect too, so we include the MSA too.

Table 8: The total number of articles that demonstrated OM of Arabic language in the listed Arabic Dialects.

Dialect	Articles count
General	30
MSA	21
Egyptian	13
Saudi	12
Algerian	9
Jordanian	8

Doi : https://doi.org/10.54216/FPA.090101
Received: April 10, 2022 Accepted: September 22, 2022

Levantine	7
Gulf	6
Moroccan	3
Lebanese	3
Arabizi	2
Syrian	2
Tunisian	2
Sudanese	1
Iraqi	1
UAE	1
North Africa	1

4.3. Dataset source

Our Articles datasets vary between tweets, Facebook comments, and post, movies review, restaurant reviews, trips reviews, other corpora, or not mentioned at all, so the Datasets in the related articles were analyzed then we categorize them into four categories:

Tweets for the datasets gathered from Twitter in certain conditions, Facebook for the dataset gathered from Facebook comments or posts, User Reviews for datasets collected from a user review in websites other than Facebook and Twitter, Other for datasets gathered from other corpus or not mentioned clearly in the article. Table 9 illustrates the dataset category and the number of articles that used it; it was clear that Twitter has the highest portion of articles; this may due not only to twitter's publicity but also it's easy to use available APIs.

Table 9. The total number of articles that demonstrated OM of the Arabic language using different datasets sources.

Dataset Source Category	Articles count
Tweets	35
Ready corpus/not mentioned	28
User reviews	20
Facebook comments/posts	16

Moreover, each method that was successful in getting the best result according to the article on these datasets groups was registered, for each dialect, we have a pie chart that shows the methods used and gave the best results according to the articles. Starting from the Twitter dataset, we can see that 35 papers were considering datasets extracted from Twitter and after making certain preprocessing such as stemming or unifying some characters or deleting duplicates. In the Twitter dataset, we can see SVM dominates other methods; second place is for lexicon-based methods as shown in Figure 2.

Doi : https://doi.org/10.54216/FPA.090101
Received: April 10, 2022 Accepted: September 22, 2022

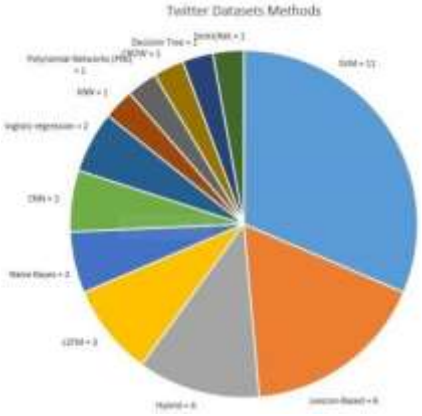

Figure 2:Twitter dataset methods.

The second top used dataset category was user reviews datasets, where data was extracted from reviews for movies, restaurants, traveling trips, books, or services. In Figure 3, we can see that SVM is also dominant as the best results were reached after applying it to most articles, below is the pie chart for the user reviews dataset and methods used on it.

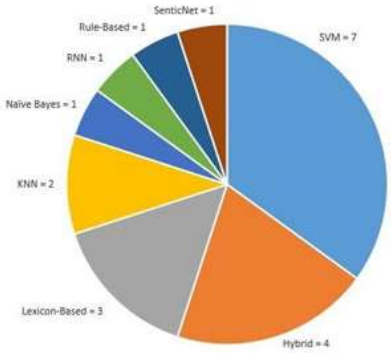

Figure 3: User review methods used.

The third-place used dataset was the ones extracted from Facebook comments or posts, here also SVM was dominant with five articles after it lexicon-based methods with four times. Figure 4 shows the distribution of the methods used.

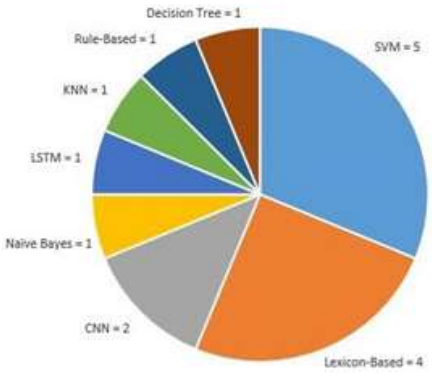

Figure 4: Facebook dataset methods.

Doi : https://doi.org/10.54216/FPA.090101
Received: April 10, 2022 Accepted: September 22, 2022

Other Datasets are the ones that are extracted from a ready corpus or the source is not mentioned clearly, so we group these articles in a single category. Methods used for this group were extracted, and Figure 5 showed the distribution of the methods used and the best ones according to the papers reviewed. In this category, lexicon-based was the dominant with 11 nomination times next was SVM with seven times, the chart below shows the all used methods.

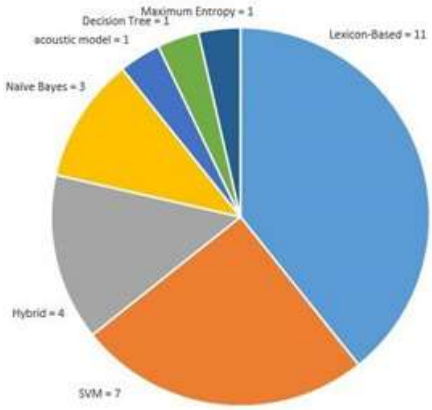

Figure 5: Facebook dataset methods.

4.4. Articles Sources

For the related articles, we have extracted the conferences and journals that published the papers, so we can nominate the top conferences and journals that publish in opinion mining for Arabic dialects from social media. Tables 10 and 11 shows the top 5 conferences and journal publishing on this subject, it illustrates that the conference international conference of women and underrepresented minorities in natural language processing (WiNLP) was the top conference publishing with five papers, and the journal Procedia Computer Science with nine articles, the top 5 lists are illustrated in the Tables 10 and 11.

Table 10:The total number of articles that demonstrated OM of Arabic language in the listed conferences.

Conference	Articles count
International conference of women and underrepresented minorities in natural language processing (WiNLP)	5
Arabic Natural Language Processing Workshop	2
International Conference on Computational Linguistics and Intelligent Text Processing	2
International conference on information and communication systems (ICICS)	2
International Conference on Language Resources and Evaluation (LREC 2016)	2

Table 11: The total number of articles that demonstrated OM of Arabic language in the listed journals.

Journal	Articles count
Procedia Computer Science	9
Journal of Cleaner Production	5
arXiv preprin arXiv:1709.08521	4

Doi : https://doi.org/10.54216/FPA.090101
Received: April 10, 2022 Accepted: September 22, 2022

International Journal of Advanced Computer Science and Applications (IJACSA)	3
ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)	2

4.5. Methods per year

For the methods used, we have sorted them by the year to show the methods trend over the four years, we have listed the 16 methods with the number of articles per year. We have noticed that Maximum Entropy appears one time, and in 2016, which can indicate it is an old method and better not to try it, the same thing can be applied to the acoustic model since it also appeared once in 2018. Hybrid methods look to be declining over the year since it did not appear in 2019, and only two articles nominate it in 2018, while it was 4,6 in 2016,2017 respectively.

On the other hand, we can see in Table 12 that rule-based is used once in 2016, and once in 2019, which may show was less effective and still less effective than other methods. For the CBOW, Polynomial Networks (PNs), and SenticNet methods even though it is low in number, it appears in 2018 and 2019 which means it is a promising future and deserves exploring and more experiments to be conducted.

Table 12: A total number of articles that demonstrated the listed methods of OM in the Arabic language from 2016 through 2019.

Method	2016	2017	2018	2019
SVM	7	9	8	6
Naive Bayes	2	1	3	1
LSTM	0	0	0	4
CNN	0	0	2	2
KNN	2	1	1	0
RNN	0	0	1	0
PNs	0	0	0	1
Rule-Based	1	0	0	1
Lexicon Based	8	11	4	1
Acoustic model	0	0	1	0
Logistic regression	0	0	0	2
CBOW	0	0	0	1
Decision Tree	1	1	0	1
Hybrid	4	6	2	0
SenticNet	0	0	2	0
Maximum Entropy	1	0	0	0

The methods that have a high number of nominations or some trends are charted in Figure 6 to show the trend over the four years. We can notice that SVM has approximately a study state and it appears in the four years with close frequencies the average for the four years is 7.5, the same thing can be said for Naıve Bayes since it appears in the four years too with close frequencies but the average is 2.75 which is much less than SVM.

In Hybrid and Lexicon-Based methods we can notice the falling trend, since they were high in 2016 and 2017 but much lower in 2018 and 2019, whereas hybrid does not appear in 2019 and lexicon-based become 2 in 2019, so we can conclude that they can be eliminated in future studied to make space for the more promising

Doi : https://doi.org/10.54216/FPA.090101
Received: April 10, 2022 Accepted: September 22, 2022

methods. The decision tree method has very low appearances, since it appeared once in 2016 and once in 2019, so it is not promising for future work.

On the other hand LSTM, CNN, and KNN appear to be promising and deserve more exploration, these methods under deep learning techniques and show better results in recent years, starting from 2016 they mostly were zero used, but after that, it starts to show increasing trend wherein 2019 it reached a peak and suppose to continue increasing.

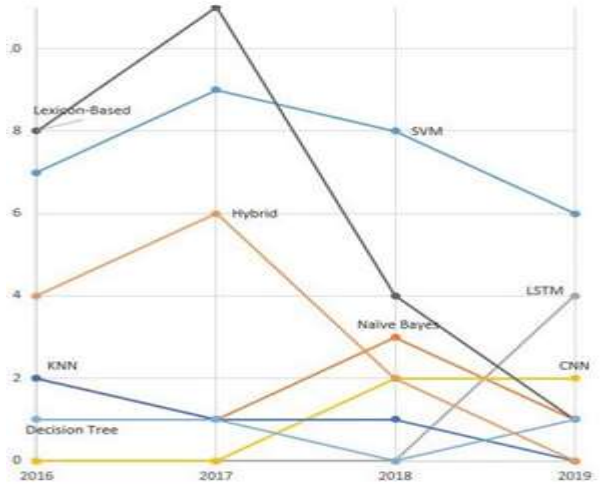

Figure 6: Frequency of different methods of OM in the Arabic language from 2016 through 2019.

4.6. Dialects and Methods

Moreover, we have analyzed each dialect and the methods that show the best results, and the number of articles that reached these results. In Table 13 we can see many dialects that have a low number of articles like the Iraqi, Syrian, and Sudanese dialects where each one is covered by only one article, this means it has a huge shortage in opinion mining efforts, and appear to be promising to study since they are used by a high number of the population who has access to the social media. And since we found that LTSM and CNN are promising, we can see that these two new methods are not used with many extensively studied dialects like Egyptian and Jordanian. North Africa and Arabizi need more study and considered with the least studied dialects, moreover, Arabizi needs more work to be translated to Arabic text than experimented with. SVM and Lexicon-Based showed the best results for most of the Arabic dialects. Many Arabic dialects are still not studied in this period, and it will be a good chance to try new techniques of them, and dialects like Mauritanian, Yemeni, Libyan, and somalin.

Table 13: Distribution of OM methods over the different dialects of the Arabic language.

Method \ Dialect	General	Jordanian	MSA	Egyptian	Saudi	Gulf	Levantine	Syrian	Moroccan	Algerian	Sudanese	Iraqi	Lebanese	Tunisian	UAE	North Africa	Arabizi
SVM	7	5	6	3	1			1	2			1	2	2		1	
Naïve Bayes	1	1	3		1												1
LSTM	1				1				1								
CNN	2								1	1							

Doi : https://doi.org/10.54216/FPA.090101
Received: April 10, 2022 Accepted: September 22, 2022

KNN	2		1								1						
RNN			1	1		1	1										
Polynomial Networks				1		1	1										
Rule-Based			1	2		1	1										
Lexicon-Based	8	1	5	3	3	2	2			5			2		1		1
Decision Tree			2														
Hybrid	5	1	2	3	1	1											
Logistic regression					1		1										
CBOW								1									

5. Conclusion

In Analyze Arabic dialect opinion mining from social media still require more and more effort, since we have systematically reviewed 223 articles covering the last four years from google scholar, after review, we nominated 103 related articles according to our inclusion and exclusion criteria, From each article dialect, dataset, and methods were extracted to analyze. We found that some Arabic dialects still don't have any effort like Mauritanian, Yemeni, Libyan, and somalin. Moreover, most countries like Egypt and Jordan have many spoken dialects which are mostly not well categorized and no special corpus for each sub-dialects, and the whole countries were studied as a single dialect. We have found four main categories' for datasets used: Twitter, Facebook, user reviews, and ready corpora, the topmost successful methods with these categories were SM and Lexicon-Based. We have studied the methods used with the year of publishing we have found that SVM is still the most successful method that produces the best results, while Lexicon-Based is decreasing over the last four years. On the other hand, we have found that deep learning methods like LSTM and CNN are become used more public and more used in the last two years. Gaps in the methods used for each dialect were pointed out many dialects are not well experimented and we have pointed out the shortage in each dialect and the methods used and several articles discussing the dialect, so in the future, we or other researchers can study it and try to fill the gap. Many challenges and difficulties were interfaced, some of them were mitigated but the others we couldn't, like the no fund issue limited the number of papers since some of the articles requires subscription charges. Search conditions can be extended to cover extra related subjects like sentiment analysis. We can expand this systematic review by covering a larger period by extending the period to cover the last 8 years instead of the last four years only. More analysis can be done on the dataset with each dialect and method used per year to show to trend and the gap for each dataset category experiment. The gaps found in certain dialects can be studied and an experiment can be conducted to build an opinion mining classifier. Dialects like Mauritanian, Yemeni, Libyan, and somalin can be studied to build a corpus and opinion mining classifier since they are never explored before.

Acknowledgments

This work is a part of a project undertaken at the British University in Dubai.

References

[1] S. A. Salloum, M. Al-Emran, A. Monem, and K. Shaalan, "A Survey of Text Mining in Social Media: Facebook and Twitter Perspectives," *Adv. Sci. Technol. Eng. Syst. J.*, vol. 2, no. 1, pp. 127–133, 2017.

[2] S. A. Salloum, M. Al-Emran, A. A. Monem, and K. Shaalan, "A survey of text mining in social media: facebook and twitter perspectives," *Adv. Sci. Technol. Eng. Syst. J*, vol. 2, no. 1, pp. 127–133, 2017.

[3] S. A. Salloum, M. Al-Emran, M. Habes, M. Alghizzawi, M. A. Ghani, and K. Shaalan, "What Impacts the Acceptance of E-learning Through Social Media? An Empirical Study," *Recent Adv. Technol. Accept. Model. Theor.*, pp. 419–431, 2021.

[4] S. Al-Skaf, E. Youssef, M. Habes, K. Alhumaid, and S. A. Salloum, "The Acceptance of Social Media Sites: An Empirical Study Using PLS-SEM and ML Approaches," in *Advanced Machine Learning Technologies and Applications: Proceedings of AMLTA 2021*, 2021, pp. 548–558.

[5] M. Habes, M. Alghizzawi, R. Khalaf, S. A. Salloum, and M. A. Ghani, "The Relationship between

Doi : https://doi.org/10.54216/FPA.090101
Received: April 10, 2022 Accepted: September 22, 2022

Social Media and Academic Performance: Facebook Perspective," *Int. J. Inf. Technol. Lang. Stud.*, vol. 2, no. 1, 2018.

[6] M. Alshurideh, S. A. Salloum, B. Al Kurdi, and M. Al-Emran, "Factors affecting the social networks acceptance: An empirical study using PLS-SEM approach," in *ACM International Conference Proceeding Series*, 2019, vol. Part F1479.

[7] A. Y. Zainal, H. Yousuf, and S. A. Salloum, "Mining social media text: extracting knowledge from Facebook," in *Joint European-US Workshop on Applications of Invariance in Computer Vision*, 2020, pp. 762–772.

[8] S. A. Salloum, C. Mhamdi, M. Al-Emran, and K. Shaalan, "Analysis and Classification of Arabic Newspapers' Facebook Pages using Text Mining Techniques," *Int. J. Inf. Technol. Lang. Stud.*, vol. 1, no. 2, pp. 8–17, 2017.

[9] M. Alghizzawi, S. A. Salloum, and M. Habes, "The role of social media in tourism marketing in Jordan," *Int. J. Inf. Technol. Lang. Stud.*, vol. 2, no. 3, 2018.

[10] M. Alghizzawi, M. Habes, S. A. Salloum, M. A. Ghani, C. Mhamdi, and K. Shaalan, "The effect of social media usage on students'e-learning acceptance in higher education: A case study from the United Arab Emirates," *Int. J. Inf. Technol. Lang. Stud.*, vol. 3, no. 3, 2019.

[11] C. Mhamdi, M. Al-Emran, and S. A. Salloum, *Text mining and analytics: A case study from news channels posts on Facebook*, vol. 740. 2018.

[12] M. Habes, S. A. Salloum, M. Alghizzawi, and C. Mhamdi, *The Relation Between Social Media and Students' Academic Performance in Jordan: YouTube Perspective*, vol. 1058. 2020.

[13] S. F. S. Alhashmi, S. A. Salloum, and C. Mhamdi, "Implementing Artificial Intelligence in the United Arab Emirates Healthcare Sector: An Extended Technology Acceptance Model," *Int. J. Inf. Technol. Lang. Stud.*, vol. 3, no. 3, 2019.

[14] S. A. Salloum and K. Shaalan, "Adoption of E-Book for University Students," in *International Conference on Advanced Intelligent Systems and Informatics*, 2018, pp. 481–494.

[15] M. T. Alshurideh, B. Al Kurdi, and S. A. Salloum, "The moderation effect of gender on accepting electronic payment technology: a study on United Arab Emirates consumers," *Rev. Int. Bus. Strateg.*, 2021.

[16] A. Aburayya *et al.*, "An Empirical Examination of the Effect of TQM Practices on Hospital Service Quality: An Assessment Study in UAE Hospitals."

[17] R. Saeed Al-Maroof, K. Alhumaid, and S. Salloum, "The Continuous Intention to Use E-Learning, from Two Different Perspectives," *Educ. Sci.*, vol. 11, no. 1, p. 6, 2020.

[18] M. Alghizzawi, M. Habes, and S. A. Salloum, *The Relationship Between Digital Media and Marketing Medical Tourism Destinations in Jordan: Facebook Perspective*, vol. 1058. 2020.

[19] R. S. Al-Maroof, S. A. Salloum, A. Q. AlHamadand, and K. Shaalan, "Understanding an Extension Technology Acceptance Model of Google Translation: A Multi-Cultural Study in United Arab Emirates," *Int. J. Interact. Mob. Technol.*, vol. 14, no. 03, pp. 157–178, 2020.

[20] R. Al-Maroof *et al.*, "The acceptance of social media video for knowledge acquisition, sharing and application: A com-parative study among YouTube users and TikTok Users' for medical purposes," *Int. J. Data Netw. Sci.*, vol. 5, no. 3, pp. 197–214, 2021.

[21] R. S. Al-Maroof, M. T. Alshurideh, S. A. Salloum, A. Q. M. AlHamad, and T. Gaber, "Acceptance of Google Meet during the spread of Coronavirus by Arab university students," in *Informatics*, 2021, vol. 8, no. 2, p. 24.

[22] Q. A. Al-Radaideh and G. Y. Al-Qudah, "Application of rough set-based feature selection for Arabic sentiment analysis," *Cognit. Comput.*, vol. 9, no. 4, pp. 436–445, 2017.

[23] S. A. Salloum, M. Al-Emran, and K. Shaalan, "A Survey of Lexical Functional Grammar in the Arabic Context," *Int. J. Com. Net. Tech*, vol. 4, no. 3, 2016.

[24] S. S. A. Al-Maroof R.S., "An Integrated Model of Continuous Intention to Use of Google Classroom.," *Al-Emran M., Shaalan K., Hassanien A. Recent Adv. Intell. Syst. Smart Appl. Stud. Syst. Decis. Control. vol 295. Springer, Cham*, 2021.

[25] "Unisco.org https://tinyurl.com/y5hllrm4."

[26] "https://tinyurl.com/yyj3yvwu."

[27] I. Guellil and F. Azouaou, "ASDA: Analyseur Syntaxique du Dialecte Alg {\'e} rien dans un but d'analyse s {\'e} mantique," *arXiv Prepr. arXiv1707.08998*, 2017.

[28] H. H. Mustafa, A. Mohamed, and D. S. Elzanfaly, "An Enhanced Approach for Arabic Sentiment Analysis," *Int. J. Artif. Intell. Appl.*, vol. 8, no. 5, 2017.

[29] A. Assiri, A. Emam, and H. Al-Dossari, "Saudi twitter corpus for sentiment analysis," *World Acad. Sci. Eng. Technol. Int. J. Comput. Electr. Autom. Control Inf. Eng.*, vol. 10, no. 2, pp. 272–275, 2016.

[30] 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 Prepr. arXiv1906.01830*, 2019.

Doi : https://doi.org/10.54216/FPA.090101
Received: April 10, 2022 Accepted: September 22, 2022

[31] A. Alawami, "Aspect Terms Extraction of Arabic Dialects for Opinion Mining Using Conditional Random Fields," in *International Conference on Intelligent Text Processing and Computational Linguistics*, 2016, pp. 211–220.
[32] A. Oussous, F.-Z. Benjelloun, A. A. Lahcen, and S. Belfkih, "ASA: A framework for Arabic sentiment analysis," *J. Inf. Sci.*, p. 0165551519849516, 2019.
[33] A. A. FE-z El-taher, "Hammouda, and S. Abdel-Mageid,'Automation of understanding textual contents in social networks,'" in *Selected Topics in Mobile & Wireless Networking (MoWNeT), 2016 International Conference on. IEEE*, 2016, pp. 1–7.
[34] G. Alwakid, T. Osman, and T. Hughes-Roberts, "Challenges in sentiment analysis for Arabic social networks," *Procedia Comput. Sci.*, vol. 117, pp. 89–100, 2017.
[35] I. Guellil, A. Adeel, F. Azouaou, F. Benali, A. Hachani, and A. Hussain, "Arabizi sentiment analysis based on transliteration and automatic corpus annotation," in *Proceedings of the 9th workshop on computational approaches to subjectivity, sentiment and social media Analysis*, 2018, pp. 335–341.
[36] K. Abainia, "DZDC12: a new multipurpose parallel Algerian Arabizi–French code-switched corpus," *Lang. Resour. Eval.*, pp. 1–37, 2019.
[37] B. Haidar, M. Chamoun, and A. Serhrouchni, "Multilingual cyberbullying detection system: Detecting cyberbullying in Arabic content," in *2017 1st Cyber Security in Networking Conference (CSNet)*, 2017, pp. 1–8.
[38] O. Al-Harbi, "A Comparative Study of Feature Selection Methods for Dialectal Arabic Sentiment Classification Using Support Vector Machine," *arXiv Prepr. arXiv1902.06242*, 2019.
[39] A. Alsayat and N. Elmitwally, "A comprehensive study for Arabic Sentiment Analysis (Challenges and Applications)," *Egypt. Informatics J.*, 2019.
[40] T. Al-Moslmi, M. Albared, A. Al-Shabi, N. Omar, and S. Abdullah, "Arabic senti-lexicon: Constructing publicly available language resources for Arabic sentiment analysis," *J. Inf. Sci.*, vol. 44, no. 3, pp. 345–362, 2018.
[41] I. Akour, N. Alnazzawi, R. Alfaisal, and S. A. Salloum, "USING CLASSICAL MACHINE LEARNING FOR PHISHING WEBSITES DETECTION FROM URLS."
[42] H. Yousuf, A. Q. Al-Hamad, and S. Salloum, "An Overview on CryptDb and Word2vec Approaches."
[43] A. Wahdan, S. A. Salloum, and K. Shaalan, "Qualitative study in Natural Language Processing: text classification," in *International Conference on Emerging Technologies and Intelligent Systems*, 2021, pp. 83–92.
[44] F. Almatrooshi, S. Alhammadi, S. A. Salloum, and K. Shaalan, "Text and web content mining: a systematic review," in *International Conference on Emerging Technologies and Intelligent Systems*, 2021, pp. 79–87.
[45] S. Salloum, T. Gaber, S. Vadera, and K. Shaalan, "Phishing Website Detection from URLs Using Classical Machine Learning ANN Model," in *International Conference on Security and Privacy in Communication Systems*, 2021, pp. 509–523.
[46] A. A. Khamees, H. D. Hejazi, M. Alshurideh, and S. A. Salloum, "Classifying Audio Music Genres Using CNN and RNN," *Adv. Mach. Learn. Technol. Appl. Proc. AMLTA 2021*, p. 315.
[47] A. Wahdan, S. A. Salloum, and K. Shaalan, "Text Classification of Arabic Text: Deep Learning in ANLP," in *Advanced Machine Learning Technologies and Applications: Proceedings of AMLTA 2021*, 2021, pp. 95–103.
[48] S. A. Salloum, "Classifying Audio Music Genres Using a Multilayer Sequential Model," *Adv. Mach. Learn. Technol. Appl. Proc. AMLTA 2021*, p. 301.
[49] S. A. Salloum, "Sentiment Analysis in Dialectal Arabic: A Systematic Review," *Adv. Mach. Learn. Technol. Appl. Proc. AMLTA 2021*, p. 407.
[50] H. Yousuf, M. Lahzi, S. A. Salloum, and K. Shaalan, "A systematic review on sequence-to-sequence learning with neural network and its models.," *Int. J. Electr. Comput. Eng.*, vol. 11, no. 3, 2021.
[51] H. Yousuf and S. Salloum, "Survey Analysis: Enhancing the Security of Vectorization by Using word2vec and CryptDB."
[52] A. Alshamsi, R. Bayari, and S. Salloum, "Sentiment analysis in English Texts," *Adv. Sci. Technol. Eng. Syst.*, vol. 5, no. 6, 2020.
[53] S. A. Salloum, M. Al-Emran, and K. Shaalan, "Mining Text in News Channels: A Case Study from Facebook," *Int. J. Inf. Technol. Lang. Stud.*, vol. 1, no. 1, pp. 1–9, 2017.
[54] S. A. Salloum, R. Khan, and K. Shaalan, "A Survey of Semantic Analysis Approaches," in *Joint European-US Workshop on Applications of Invariance in Computer Vision*, 2020, pp. 61–70.
[55] S. A. Salloum, M. Al-Emran, A. A. Monem, and K. Shaalan, *Using text mining techniques for extracting information from research articles*, vol. 740. 2018.
[56] M. Alkhatib, M. El Barachi, and K. Shaalan, "An Arabic social media based framework for incidents

Doi : https://doi.org/10.54216/FPA.090101
Received: April 10, 2022 Accepted: September 22, 2022

and events monitoring in smart cities," *J. Clean. Prod.*, vol. 220, pp. 771–785, 2019.
[57] H. G. Hassan, H. M. A. Bakr, and B. E. Ziedan, "A framework for Arabic concept-level sentiment analysis using SenticNet," *Int. J. Electr. Comput. Eng.*, vol. 8, no. 5, p. 4015, 2018.
[58] M. S. Al-Batah, S. Mrayyen, and M. Alzaqebah, "Investigation of Naive Bayes Combined with Multilayer Perceptron for Arabic Sentiment Analysis and Opinion Mining.," *JCS*, vol. 14, no. 8, pp. 1104–1114, 2018.
[59] A. Alqarafi, A. Adeel, A. Hawalah, K. Swingler, and A. Hussain, "A Semi-supervised Corpus Annotation for Saudi Sentiment Analysis Using Twitter," in *International Conference on Brain Inspired Cognitive Systems*, 2018, pp. 589–596.
[60] W. Adouane and R. Johansson, "Gulf Arabic Linguistic Resource Building for Sentiment Analysis," in *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, 2016, pp. 2710–2715.
[61] O. Al-Harbi, "Classifying Sentiment of Dialectal Arabic Reviews: A Semi-Supervised Approach," *Int. Arab J. Inf. Technol.*, vol. 16, no. 6, pp. 995–1002, 2019.
[62] H. Najadat, A. Al-Abdi, and Y. Sayaheen, "Model-based sentiment analysis of customer satisfaction for the Jordanian telecommunication companies," in *2018 9th International Conference on Information and Communication Systems (ICICS)*, 2018, pp. 233–237.
[63] A. Elouardighi, M. Maghfour, and H. Hammia, "Collecting and processing arabic facebook comments for sentiment analysis," in *International Conference on Model and Data Engineering*, 2017, pp. 262–274.
[64] R. M. Alahmary, H. Z. Al-Dossari, and A. Z. Emam, "Sentiment Analysis of Saudi Dialect Using Deep Learning Techniques," in *2019 International Conference on Electronics, Information, and Communication (ICEIC)*, 2019, pp. 1–6.
[65] R. Ismail, M. Omer, M. Tabir, N. Mahadi, and I. Amin, "Sentiment Analysis for Arabic Dialect Using Supervised Learning," in *2018 International Conference on Computer, Control, Electrical, and Electronics Engineering (ICCCEEE)*, 2018, pp. 1–6.
[66] M. N. Al-Kabi, H. A. Wahsheh, and I. M. Alsmadi, "Polarity classification of Arabic sentiments," *Int. J. Inf. Technol. Web Eng.*, vol. 11, no. 3, pp. 32–49, 2016.
[67] B. Haidar, M. Chamoun, and A. Serhrouchni, "A multilingual system for cyberbullying detection: Arabic content detection using machine learning," *Adv. Sci. Technol. Eng. Syst. J*, vol. 2, no. 6, pp. 275–284, 2017.
[68] N. Al-Twairesh *et al.*, "Suar: Towards building a corpus for the Saudi dialect," *Procedia Comput. Sci.*, vol. 142, pp. 72–82, 2018.
[69] H. K. Aldayel and A. M. Azmi, "Arabic tweets sentiment analysis–a hybrid scheme," *J. Inf. Sci.*, vol. 42, no. 6, pp. 782–797, 2016.
[70] A. A. A. Mehrez, M. Alshurideh, B. A. Kurdi, and S. A. Salloum, *Internal Factors Affect Knowledge Management and Firm Performance: A Systematic Review*, vol. 1261 AISC. 2021.
[71] F. Al Suwaidi, M. Alshurideh, B. Al Kurdi, and S. A. Salloum, *The Impact of Innovation Management in SMEs Performance: A Systematic Review*, vol. 1261 AISC. 2021.
[72] H. AlShehhi, M. Alshurideh, B. A. Kurdi, and S. A. Salloum, *The Impact of Ethical Leadership on Employees Performance: A Systematic Review*, vol. 1261 AISC. 2021.
[73] A. Ahmed, M. Alshurideh, B. Al Kurdi, and S. A. Salloum, *Digital Transformation and Organizational Operational Decision Making: A Systematic Review*, vol. 1261 AISC. 2021.
[74] A. Alshamsi, M. Alshurideh, B. A. Kurdi, and S. A. Salloum, *The Influence of Service Quality on Customer Retention: A Systematic Review in the Higher Education*, vol. 1261 AISC. 2021.
[75] R. Al-Maroof, N. Al-Qaysi, S. A. Salloum, and M. Al-Emran, "Blended Learning Acceptance: A Systematic Review of Information Systems Models," *Technol. Knowl. Learn.*, pp. 1–36, 2021.
[76] S. F. S. Alhashmi *et al.*, "A Systematic Review of the Factors Affecting the Artificial Intelligence Implementation in the Health Care Sector.," in *AICV*, 2020, pp. 37–49.
[77] K. S. A. Wahdan, S. Hantoobi, S. A. Salloum, and K. Shaalan, "A systematic review of text classification research based ondeep learning models in Arabic language," *Int. J. Electr. Comput. Eng*, vol. 10, no. 6, pp. 6629–6643, 2020.
[78] S. K. Areed S., Salloum S.A., "The Role of Knowledge Management Processes for Enhancing and Supporting Innovative Organizations: A Systematic Review.," *Al-Emran M., Shaalan K., Hassanien A. Recent Adv. Intell. Syst. Smart Appl. Stud. Syst. Decis. Control. vol 295. Springer, Cham*, 2021.
[79] S. K. Al Mansoori S., Salloum S.A., "The Impact of Artificial Intelligence and Information Technologies on the Efficiency of Knowledge Management at Modern Organizations: A Systematic Review.," *Al-Emran M., Shaalan K., Hassanien A. Recent Adv. Intell. Syst. Smart Appl. Stud. Syst. Decis. Control. vol 295. Springer, Cham*, 2021.
[80] S. K. Yousuf H., Lahzi M., Salloum S.A., "Systematic Review on Fully Homomorphic Encryption

Doi : https://doi.org/10.54216/FPA.090101
Received: April 10, 2022 Accepted: September 22, 2022

Scheme and Its Application.," *Al-Emran M., Shaalan K., Hassanien A. Recent Adv. Intell. Syst. Smart Appl. Stud. Syst. Decis. Control. vol 295. Springer, Cham*, 2021.

[81] R. ALBayari, S. Abdullah, and S. A. Salloum, "Cyberbullying Classification Methods for Arabic: A Systematic Review," in *The International Conference on Artificial Intelligence and Computer Vision*, 2021, pp. 375–385.

[82] S. Hantoobi, A. Wahdan, S. A. Salloum, and K. Shaalan, "Integration of Knowledge Management in a Virtual Learning Environment: A Systematic Review," *Recent Adv. Technol. Accept. Model. Theor.*, pp. 247–272, 2021.

[83] S. Kitchenham, B, Charters, "Guidelines for performing systematic literature reviews in software engineering.," *Softw. Eng. Group, Sch. Comput. Sci. Math. Keele Univ. 1–57.*, 2007.

Doi : https://doi.org/10.54216/FPA.090101
Received: April 10, 2022 Accepted: September 22, 2022