

Classification of Arabic Tweets: A Review

Meshrif Alruily

Faculty of Computer and Information Sciences, Jouf University, Sakaka 72388, Saudi Arabia; mfalruily@ju.edu.sa

Abstract: Text classification is a prominent research area, gaining more interest in academia, industry and social media. Arabic is one of the world's most famous languages and it had a significant role in science, mathematics and philosophy in Europe in the middle ages. During the Arab Spring, social media, that is, Facebook, Twitter and Instagram, played an essential role in establishing, running, and spreading these movements. Arabic Sentiment Analysis (ASA) and Arabic Text Classification (ATC) for these social media tools are hot topics, aiming to obtain valuable Arabic text insights. Although some surveys are available on this topic, the studies and research on Arabic Tweets need to be classified on the basis of machine learning algorithms. Machine learning algorithms and lexicon-based classifications are considered essential tools for text processing. In this paper, a comparison of previous surveys is presented, elaborating the need for a comprehensive study on Arabic Tweets. Research studies are classified according to machine learning algorithms, supervised learning, unsupervised learning, hybrid, and lexicon-based classifications, and their advantages/disadvantages are discussed comprehensively. We pose different challenges and future research directions.

Keywords: Arabic language processing; Arabic sentiment analysis; Twitter data analysis; natural language processing



Citation: Alruily, M. Classification of Arabic Tweets: A Review. *Electronics* **2021**, *10*, 1143. <https://doi.org/10.3390/electronics10101143>

Academic Editors: Patricia Martín-Rodilla and Pablo Gamallo

Received: 17 March 2021

Accepted: 5 May 2021

Published: 12 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

A blend of three elements refers to social media: data, user groups and Web 2.0 technologies [1]. It has continued to grow since the launch of the first social media networks about two decades ago, and now it is an integral part of people's daily lives [2,3]. By using social media, we can share news, opinions and communicate with anyone worldwide [4]. Social media platforms such as Twitter can be used as a valuable tool for sentiment analysis as people use it to share their viewpoints on a wide variety of topics. The quantity and quality of data from social media are rising dramatically [5]. On average, Twitter has 500 million tweets daily generated by more than 230 million active users. [6]. Among all social media platforms, Twitter is one of the best and useful platforms for information sharing. Twitter has 330 million active users per month with a beneficial module called a hashtag representing a group conversation on Twitter, and some users may use hashtags to express their opinions about any matter [7]. In 2011, during the Arab Spring, the most prominent Twitter hashtags throughout the Arab world were #Jan25, #Libya, #Brain, and #protests. In the first three months of the year, there were 1.4 million of hashtags for "# Egypt" and 1.2 million for # Jan25. Twitter's main strength is not its number of users nor the huge number of tweets posted on it but the fact that companies and political parties are well aware that many of their clients and supporters are on Twitter. They can easily communicate with them at once [8].

Sentiment analysis is a valuable approach for gaining insights from a massive number of tweets shared by multiple users [9,10]. We can classify the sentiments of people into positive, negative, or neutral categories. The approaches used for Arabic text classification can be divided into two main categories, that is, machine learning techniques and semantic orientation techniques. In these techniques, the data is gathered to build the data set to train the machine learning model [11]. After training, the model is used to predict the given

tweet/data [12]. In the second technique, sentiment lexicons of the language are created. Every word in the lexicon is given a degree of positivity and negativity. This degree of the words indicates their class, that is, positive, negative and neutral. Any document classified into these classes based on the sum of degrees of all words used in the document [13]. In the last decade, sentiment analysis has gained the attention of researchers. It is a top trending research area in the current era. The application of sentiment analysis is everywhere and increasing day by day. A lot of tools and techniques are developed by researchers for the analysis of various languages. However, a little work is done on the Arabic language, especially on Arabic Tweets. According to a survey, more than 300 million speakers of the Arabic language [13].

Most of the work done on sentiment analysis till now is on classifying sentiments in text into three distinct categories. These categories are positive, negative, and neutral. This work is not enough for making valuable decisions as it may cause ambiguous results [14]. For instance, if a sentence is classified as a positive or negative sentence, it is unclear when the text's polarity is what negative or positive emotion the user has. Every sentence can have sub-emotion categories. A sentence can have both positive and negative emotions. Moreover, sarcasm is also a critical factor in sentiment analysis. The detection of sarcasm is another research issue in language processing. The degree of sentiment is also significant in the classification of sentences. The comments posted by more than one individual may have a common sentiment, with different intensity levels [15]. Identification of a particular type of sentiment and its intensity can help analyze user opinion that can lead to a finer result/response.

In this paper, we laid a comparison of different review studies on Arabic text processing. These studies included general Arabic text processing. To the best of our knowledge, there is no detailed analysis of Arabic Tweets. We discuss the Arabic Tweet processing and relevant machine learning algorithms. For sentiment analysis of Arabic Tweets, machine learning algorithms are divided into three categories: supervised learning, unsupervised learning, hybrid techniques and lexicon-based classifications. We compare and contrast all the approaches in the form of tables, and a discussion/learned lesson section is also considered. Finally, current research challenges and future directions of Arabic Tweets are also posed. The rest of the paper is organized as follows: Section 2 deliberately discusses the review papers on Arabic Text classification. Section 3 explores various phases and techniques of machine learning. Section 4 summarizes differently supervised, unsupervised and hybrid techniques proposed for Arabic Tweets. Section 5 discusses various lexicon techniques and compares them. Section 6 explores challenges and limitations for Arabic text classification. In Section 7, deep learning for Arabic sentiment analysis is presented. Transformer for Arabic text is provided in Section 8. Section 9 presents future research directions, while Section 10 concludes the survey.

2. Comparison with Other Surveys

There are several basic research studies [16–19] on Arabic data analysis. In these studies, the authors discuss the basic knowledge and need of Arabic Sentiment Analysis (SA). The followings are the more enhanced review of Arabic data classifications and research challenges. Alhumoud et al. [20] present a study of research efforts to examine Arabic content on Twitter, with a particular emphasis on the techniques and methods used to obtain sentiments for Arabic content on Twitter. Assiri et al. [21] discuss the important studies related to the Arabic sentiment analysis and provide in-depth qualitative analysis based on various Arabic text features. They evaluate the smoothness of percentage errors of different research studies that describe the influence of these studies.

Deep Learning (DL) is becoming very popular for different language processing and online social networks (OSN). In [22], DL-based techniques used for natural language processing (NLP) and speech processing are surveyed. Most of the studies focused on Optical Character Recognition (OCR) based problems for text translations. Similarly, Abdullah et al. [16] surveyed Arabic Twitter data for sentiment analysis. They discussed

different machine learning techniques and NLP classifications for Arabic data. Different aspects and challenges of Arabic languages are discussed in detail. But it does not consider Machine Learning (ML) classifications for Arabic data. Guellil et al. [23] presented a comprehensive survey on Arabic varieties, that is, classical Arabic, Modern Standard Arabic, Arabic Dialect. Arabizi (roman script) is considered in any application nowadays. They consider a large number of research studies by categorizing them according to these varieties. Different machine learning algorithms, that is, Support Vector Machine (SVM), convolutional neural network (CNN), Naive Base (NB), for Arabic language processing.

Badaro et al. [24] present comprehensive overview of the literature on Arabic opinion mining. The literature covers all elements involved in a systematic opinion-mining system: methods, resources, approaches, and implementations on Arabic sentiment analysis. The most recent developments in the field, including cutting-edge deep learning models for sentiment analysis are discussed. An overview of a system architecture for an Arabic opinion-mining system to help people who want to build a real opinion-mining application is also elaborated. The survey [25] discusses several machine learning approaches, that is, supervised, unsupervised, and hybrid studies for Arabic text processing. Some important improvement is discussed related to SVM, N-gram studies, and NB. This survey shows that more custom algorithms and methods need to be developed for Arabic data analysis. Kaseb et al. [26] provide a comparison and study of the various proposed Arabic SA methodologies. Some ML techniques, that is, SVM, NB, are also considered for comparison purposes. In the case of SA, it was discovered that SVM produces the best results. The issue with NB is that it is focused on probabilities, so it is better suited to high-dimensional inputs. SVM has been successfully used in many sentiment analysis tasks due to its main advantages.

El-Masri et al. [27] discuss the advantages and disadvantages of different deep learning models for Arabic opinion-mining. Some other, more relevant studies [28,29] also discuss the research trends and challenges for Arabic data classifications. But all these studies lack the classification on the basis of machine learning algorithms in a comprehensive way. So we adopted different classifications of machine learning and hybrid approaches for this topic. Ghallab et al. [30] presented different classification methods for Arabic text processing. This paper highlights the most popular preprocessing techniques and feature selection methods. It also includes a taxonomy of sentiment classification techniques.

Abo et al. [31] studied different Arabic text processing approaches based on machine learning and Lexicon-based classifications. They discuss two case studies related to Arabic text. They are missing relevant discussions and learned lessons on these techniques. Alsayat et al. [32] discuss six levels of the Arabic language that can be derived based on morphology, phonetics, lexicology, syntax. They compare some techniques by using a limited number of factors. Table 1 compare different existing review papers based on different parameters. Although there are several surveys on different Arabic Sentiment Analysis topics, there is no comprehensive survey on Arabic tweets. This survey provides comprehensive and up-to-date analyses of current studies and development on this topic due to the importance of social media and Twitter-like platforms.

Table 1. Comparison of surveys on Arabic text processing.

Ref.	Year	Objectives	Techniques	Machine Learning	Arabic Tweets
Alhumoud et al. [20]	2015	Tools and techniques for Arabic data analysis	Discuss different Machine learning approaches used for SA	✓✓	NA
Al-Ayyoub et al. [22]	2019	Comprehensive analysis of general Arabic sentiment analysis	Aspect-Based SA, Binary SA, Turnary SA, Multi-Way SA, Aspect-Based SA, Multilingual SA	✓✓	Partially
Guellil et al. [23]	2019	Surveyed different Arabic varieties i.e., classical Arabic, Modern Standard Arabic, Arabic Dialect	text mining and machine learning techniques.	✓✓	Partially
Badaro et al. [24]	2019	Different tools, resources, and techniques, for Arabic text analysis	Machine learning and others	✓✓	Partially
El-Masri et al. [27]	2017	Presenting Deep learning techniques used in different applications of Arabic SA	Deep learning	Partially	Partially
Ghallab et al. [30]	2020	Journal and conference based classification	ML, DL, hybrid	✓✓	Partially
Abo et al. [31]	2020	Survey on Arabic text processing	ML, hybrid, lexicon	✓✓	Partially
This survey	2021	Classifications of Arabic Tweets	Machine Learning, hybrid	✓✓	✓✓

3. Background Knowledge

In this section, Arabic language, Arabic text classification, data gathering, and other basic concepts are discussed. Different classifications of machine learning algorithms are also elaborated.

3.1. Arabic Language

Arabic is a language ranked in the top five major languages of the world. It is commonly used in the Muslim world as the script of the Qur'an, the holy book of Islam is in Arabic. It belongs to the Semitic group of languages, comprised of Hebrew and Amharic, Ethiopia's main language. Arabic has various dialects, that is, classical, modern standard, and various local dialects. There are 28 alphabets used in this language and sentences are written from right to left. The alphabets of the Arabic language are shown in Figure 1.

The preprocessing of the Arabic language is not the same as English because of its complex and rich morphology. We need a special kind of preprocessing techniques to implement Machine and deep learning techniques for classifying Arabic tweets. Text translation, tokenization, stop-word elimination, and stemming are common activities in the preprocessing step. Stemming is the process of removing both affixes and suffixes from a word to

isolate the root. Since the Arabic language has various ways of representing text, three stemming strategies are widely used. These techniques are Khoja stemming, light stemming methods, and raw text comparison (no stemming). The feature extraction/selection role comes next. The impact of text preprocessing functions on text categorization is calculated in this step, particularly the impact of using stemming from Arabic text categorization. In this process, the term weigh is used to describe each text as a weight vector. This is commonly referred to as the bag of words process. The term frequency (tf) counts the number of times the term t appears in document d , while document frequency (df) counts the number of times the term t appears in at least one document. The inverse document frequency (IDF), on the other hand, tests how popular the word is in all records. The IDF would be low if the term exists in a large number of documents and heavy if the term appears in a small number of documents.

ARABIC ALPHABETS

kh-aa خ	H-aa ح	j-eem ج	th-aa ث	t-aa ت	b-aa ب	a-lif ا
S-aad ص	sh-een ش	s-een س	z-aa ز	r-aa ر	dh-aal ذ	d-aal د
q-aaf ق	f-aa ف	gh-ain غ	3-ain ع	Dh-aa ظ	T-aa ط	D-aad ض
y-aa ي	w-aaw و	h-a ه	n-oon ن	m-eem م	l-aam ل	k-aaf ك

Figure 1. Alphabets of Arabic Text.

3.2. Arabic Dialect

As in several languages, modern Standard Arabic (MSA) is the published and scientific standard for Arabic as this is not the word people speak in their daily lives [33]. Arabic dialects and MSA, for example, certain variations are speaking Arabic:

- Has a more basic grammar and informal language and style
- Has several distinctly articulated letters that may vary on the basis of dialect
- Has terms or phrases that differ from some dialects
- Only in writing if an intimate or humorous touch is needed

In addition to these variations between MSA and Arabic dialects, several dialects may vary [34]. As in many languages, these dialect gaps are often not large enough to make it difficult for native speakers to understand each other. As a language student, this is crucial to know, because it is more probable that you would experience a variation in dialect during conversation with native speakers. There are various types of Arabic in the Arab world. The primary variant classes include:

3.2.1. Sudanese Arabic

These dialects of Arabic are mainly spoken in Sudan and various regions of Eritrea. Sudanese Arabic is similar to the Egyptian dialect in terms of native speakers, although there are some differences [35]. More than 17 million speakers speak it. Sudanese is distinguished from other Arabic dialects by the continuation of ancient “pronunciations and writing sequences”.

3.2.2. Egyptian Arabic

With over 60 million speakers, Egyptian Arabic is the most learned and commonly spoken Arabic dialect. Egyptian Arabic influences European languages, that is, Italian, French, Greek, and Turkish.

3.2.3. Maghrebi Arabic

Maghrebi Arabic is a widely used dialect of Arabic spoken by over 70 million people worldwide. This dialect is used in various countries, that is, Algeria, Tunisia, Morocco, Western Sahara, Libya and Mauritania [36]. This dialect of Arabic is very different from modern Standard Arabic (MSA) and it has its own name that is Derja, Derija or Darija (الدارجة). This dialect is the combination of various Arabic dialects such as Algerian Arabic, Tunisian Arabic and Moroccan Arabic.

3.2.4. Gulf

According to a survey, 36 million people speak Gulf Arabic. Persian Gulf countries mostly speak it. These countries include Bahrain, Qatar, Kuwait, United Arab Emirates (UAE), Iraq, and Oman. This dialect is also made up of various dialects that are distinct in vocabulary, syntax, and pronunciation across the regions.

3.2.5. Levantine

Levantine is a dialect of Arabic spoken in Jordan, Lebanon, Palestine, and Syria and is spoken by more than 20 million speakers. It is also the second most commonly used dialect in the Arabic media. This dialect is closely related to MSA, but it has its vocabulary, phonology, and syntax.

3.2.6. Yemeni Arabic

Yemeni Arabic is spoken by nearly 15 million people who speak it as their first language. Saudi Arabia, Somalia, and Djibouti are among the countries where it is widely spoken. This dialect is not used for writing purposes and speakers of this dialect use MSA for writing purposes.

3.2.7. Mesopotamian

Mesopotamian dialect, also known as “Iraqi Arabic”, is a dialect spoken in various countries, that is, Syria, Iraq, Iran, and Turkey. It is thought to have evolved during the historical change from Aramaic to Arabic. It is spoken by about 15 million people worldwide. Many other dialects have similarities with other languages such as Akkadian, Persian, and Turkish in this Arabic dialect.

3.3. Text Classification

Text classification is known as the method of labeling data with a specific tag or label for dividing it into different categories. Text classification is one of the essential tasks in natural language processing. It has various applications like sentiment analysis, spam filtering, topic labeling, and intent identification [37]. Mostly text classification or text mining is done using supervised machine learning techniques.

Various social media platforms like Facebook, Twitter, Instagram, emails, web pages, and surveys can be used to get data. But this data is unstructured and has many ambiguities [15]. Text can be a vibrant set of information, but due to its amorphous nature, obtaining insights from it can be difficult and time-consuming. Text classification for machine learning can help organizations organize and interpret their text dynamically, quickly, and cost-effectively, simplify processes, and boost data-driven decisions. Language processing has four essential levels [38]. These four stages include data gathering, preprocessing, model training, model testing, and model deployment. These stages are discussed below.

3.4. Data Gathering

Data gathering is the first step in language processing, and we collect the desired data from various platforms like Twitter, Blogs, and other websites. Because of the evolution of the internet, the amount of data is increasing day by day. The data on internet is unstructured data and does not have any specific shape [39]. We gather data by using various techniques like web scraping and API's. It is important to gather data from reliable

resources like Facebook, Twitter, or Product data. If collected data does not have sufficient information, the results will not be trustworthy. In sentiment analysis, the data can be gathered from social media platforms like Twitter, Facebook, Instagram, and other websites like blogs and E-commerce websites [40]. Researchers used various social media platforms for data gathering in Arabic sentiment analysis, but most of the researchers used the Twitter data set as it is considered a credible source of information [41]. There is a significantly lower amount of data on Twitter that is fake [42].

3.5. Arabic Corpora

Various Arabic corpora are being used for Arabic SA, including the Quranic Arabic Corpus [43], arabiCorpus [44], Tunisian Arabic Corpus [45], International Corpus of Arabic [46], King Abdulaziz City for Science and Technology (KACST) Arabic Corpus, KALIMAT, and Arabic Corpus. These corpora have different sizes and dialects. The Quranic Arabic Corpus is an annotated corpus that includes each word in the Holy Quran's with its Arabic grammar, syntax and morphology. Three types of analysis are available in this corpus, that is, morphological annotation, syntactic treebank, and a semantic ontology. The arbi corpus is another Arabic corpus containing word frequency data and the ability to search for greater patterns and grammatical structures. Words in Arabic and Latin scripts can be searched in this corpus. Tunisian Arabic corpus is available online for free. It comprises 818,310 words that are divided into 17 categories. These categories include blogs, phone conversations, Internet forums, jokes, and so on. A researcher has three options for searching for any word in this corpus, that is, exact, stem, and regEx (transliteration). KALIMAT is a natural language Arabic resource with 6 categories that contain 4,000,000 words. These categories include history, economy, local news, international news, religion, and sport. KACST is also a freely available Arabic corpus that can be used for various research projects, including natural language processing (NLP). It has over one billion Arabic terms in it. It also covers written texts in Classical Arabic and Modern Standard Arabic (MSA) from pre-Islamic times until the corpus launch. Moreover, some other Arabic corpora are constructed by various researchers and used for various NLP tasks of Arabic language.

3.6. Exploring/Prepossessing Data

The gathered data are examined in this process, and data insights are retrieved. Since the data gathered from the internet are unstructured, it is important to preprocess it before using it in later stages. If data are not pre-processed, it can cause more computation, and incorrect results [42]. In this step, different parts of data that are not necessary are removed, like stop words, weblinks, white spaces, and special characters. After successful data pre-processing, various features are collected based on the problem of the research. In the processing of text, we represent the text in discrete and categorical values [47]. This is essential because the machine learning model is unable to comprehend textual details. Feature extraction is the method of mapping words to real-valued vectors. Various techniques can be used for feature extraction like a bag of words and TF-IDF [48]. Because of the preprocessing constraints discussed above, Arabic Natural Language Processing (NLP) tasks, such as Sentiment Classification and Named Entity Recognition (NER), are difficult to perform. Recently, there has been an increase in the use of transformers for these purposes. Language-specific BERT-based approaches have shown higher accuracy and effectiveness as they are trained on a very large corpus. There are various transformer models used for ATC, that is, AraBERT, ARBERT, MARBERT.

3.7. Train, and Evaluate Model

After the successful feature extraction, we choose a machine learning model for training. There are various machine learning models like Supervised learning models, unsupervised learning models. Models are chosen according to the data set. There is no model that we can say the best model. Different models behave contrarily on various data sets [48]. The data set is divided into two sets, that is, the training data set and the testing

data set. The selected model is trained on the training data set and evaluated on the testing data set. Various machine learning models can be trained and evaluated on the data sets to compare the accuracy and other evaluation metrics.

3.8. Deployment of Model

The last step is deploying the model for using it on real data to make use of the trained model for practical decision making. It is integrating the trained and tested machine learning model into a functional environment [49]. To get the most value from the model, it is necessary to seamlessly deploy it to production to use it to make practical decisions. Figure 2 shows the stages of text classification or language processing.

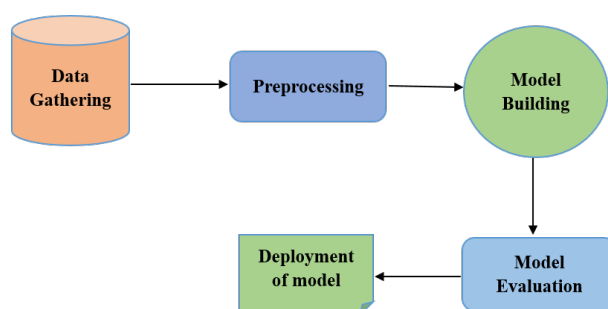


Figure 2. Phases of Text Classification.

Text classification can be done by using various techniques. These techniques are divided into three main categories, that is, machine learning, lexicon based and hybrid approaches as shown in Figure 3.

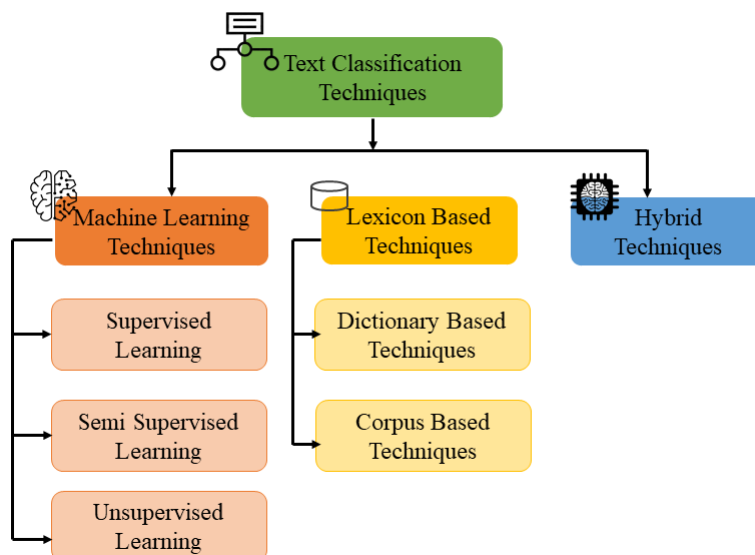


Figure 3. Techniques for Text Classification.

3.9. Machine Learning Algorithms

From a group of powerful learning techniques of AI, the ML [50] approach is extensively used for data mining, computational linguistics, and so forth the stand-alone system capable of learning from the training data [51]. It is a mechanism that can learn from the experience in the absence of explicit programming to solve the several problems, healthcare, manufacturing, text analysis, [52]. The approach of ML relies on two main phases; one is the training phase and the other one is the decision phase, as shown in Figure 4. In the first phase, the machine gets the training data and train the model according to data, and in the second phase, the system predicts the results and update itself [53,54].

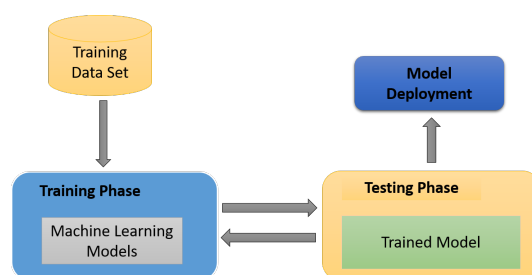


Figure 4. Process of Machine Learning.

The ML is used to develop such systems that can automatically learn from the data and can get to know about the hidden pattern. ML algorithms are grouped by their learning methodology and functional similarity in the way of working [55]. Different problem categories get benefits from ML such as classification, regression, clustering, and rule extraction [56]. The most important advantage of ML is that they deal with a complex problem and gives close or better results than a human being. In Arabic text processing, complex problems exist and need efficient solutions to solve these problems [57]. ML approaches are divided into three main subcategories for text processing as described in Figure 5.

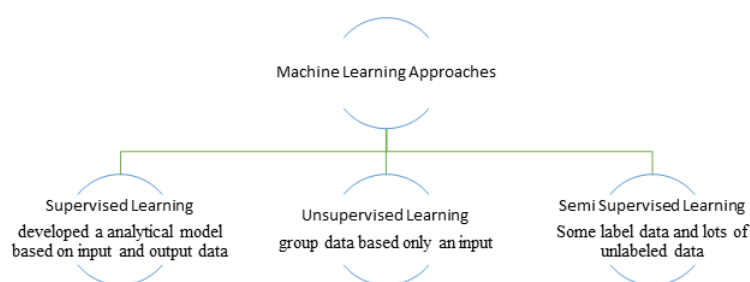


Figure 5. Machine Learning Classifications.

3.9.1. Supervised Learning

In supervised learning, the model is trained using labeled data [58]. After training the model, unknown data is provided to the system to get the expected output [59]. In this method, features are learned from the labeled input data: features are the multiple important characteristics obtained from the model and stored in a feature vector. Features are extracted from data and, together with labels, are provided to train the model. Once the model is trained, it is validated by providing unlabeled data, with the model predicting the expected label [60].

3.9.2. Unsupervised Learning

In unsupervised learning, the model provides data with no labels. The aim is to find certain patterns and knowledge in the data by creating clusters or grouping similar items together [61]. When datasets are too large to be labeled, this is the most widely-adopted approach [62]. In this method, relevant features from the data are extracted and fed to the model [63]. When the new data is given, the basic goal is to isolate groups with identical characteristics and classify them into categories.

3.9.3. Semi-Supervised Learning

Traditionally, learning has been studied either when all the data is labeled (i.e., supervised learning) or all of the information is unlabeled (i.e., unsupervised learning) [64]. In semi-supervised learning, the model is trained using both labeled and unlabeled data [65,66]. Semi-supervised learning is useful when there is a huge amount of data, and labeling all data is not possible, but it is still possible to label part of it. The goal is to

improve the system's learning behavior and performance using a combination of labeled and unlabeled data [67].

4. Machine Learning Techniques for Arabic Tweet Classification

The process of training a computer to make accurate results when any data is given to it is called machine learning. It is a type of artificial intelligence in which we make a computer system so intelligent that it can perform various tasks without human intervention [68]. Today machine learning is everywhere, from the classification of an image to self-driving cars. Different machine learning techniques can be used for language processing. In recent years most of the work done on Arabic tweet classification or Arabic sentiment analysis is done by using the following machine learning techniques [69].

4.1. Supervised Learning Techniques

Supervised learning techniques are machine learning techniques that need labeled data for building models. In this type of learning, the correct labels of data are given, and models are trained with proper tags, after training the model to predict or classify the new data based on labels given on training phase [70]. Figure 6 illustrates the process of supervised machine learning.

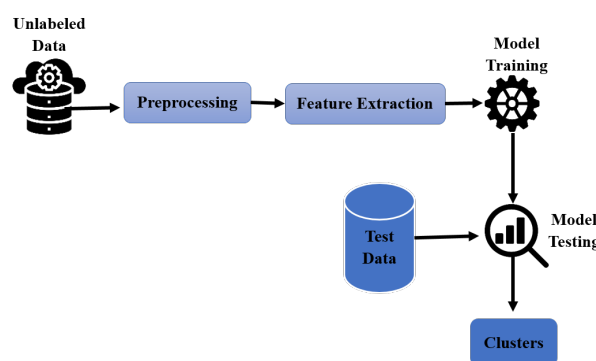


Figure 6. Structure of supervised Machine Learning.

A supervised learning model can learn to identify, provided suitable examples, the clusters of pixels and shapes associated with each number and finally recognize handwritten digits that can reliably distinguish between numbers 6 and 9 or 26 and 29. In the following section, supervised learning models for Arabic Language are explored.

Duwari et al. [71] proposed a technique for the Arabic language sentiment analysis. In this study data set of 2591 tweets were used that was collected by crowd sourcing. After the collection, tweets are labeled as positive or negative and neutral according to their properties. Three machine learning algorithms named NB, KNN, and SVM were used to classify tweets. For accurate results, 10-fold cross-validation was used. After the evaluation, the SVM classifier outperforms other classifiers by getting an accuracy of 75.25%.

Atoum et al. [72] proposed a model for tweet classification. This study also classifies tweets into three major categories, that is, positive, negative, and neutral tweets. In the data collection step, Arabic Jordanian dialect tweets were collected and preprocessed. Various tasks like text normalization, tokenization, stop words removal, name entity recognition, and stemming is done in this stage. Numerous experiments were performed on the proposed system incorporating supervised machine learning. The authors conclude that classifications using the SVM on Arabic light stemming outperform the NB classifier. Moreover, by introducing a correlation between the three categories and decreases the amount of examples to some of the most encountered instances, the accuracy of models was improved and the final accuracy of SVM was 82.1%. Jardaneh et al. [73] proposed a supervised machine learning technique for Arabic language processing. A machine learning model for quantifying the credibility of Arabic language tweets was presented in this paper. They built various features for this and divided them into two categories:

content-based features and user-based features. Different machine learning algorithms are implemented during the learning process, and the ones with the best performance are maintained. Finally, they compared each experiment's performance and presented the outcomes of each one. The experimental assessment shows that with an accuracy of 76%, the system can filter out non-credible tweets.

Al-Horaibi et al. [74] have suggested a methodology opinion mining from Arabic social media posts. In their work they use various features for sentiment classification. The twitter data set containing 2000 tweets was used in this study. The data set was manually annotated by native Arabic speakers before preprocessing for getting better accuracy. After preprocessing two supervised models that is, NB and Decision tree were applied on training data set by using different combinations of preprocessing functions. The NB model got 64.84% accuracy while decision tree got 53.75% accuracy. Various reasons of low accuracy of proposed techniques on Arabic language were also discussed in this article.

An ensemble technique was proposed by Abdelaal et al. [75] for sentiment analysis of Arabic tweets. Ensemble methods, that is, bagging, stacking and boosting models were used for enhancing the classification accuracy of Arabic text. In this study, data was gathered using twitter API and manually classified into five different categories. The categories were sports, politics, culture, technology and general. After preprocessing of data set three classifiers decision tree, NB and Sequential Minimal Optimization (SMO) were applied for the classification of tweets. After cross validation the accuracy of classifiers was 87%, 83.6%, and 86.4% respectively. Then bagging, boosting and stacking techniques were applied on models to enhance their accuracy's. After evaluation of ensemble methods, the NB model got 88.6% accuracy using bagging approach while SMO got 88.6% accuracy using stacking technique.

Alsanad et al. [76] presented an article in which they used corpus based method for the classification of Arabic tweets. They classified tweets into three distinct categories. After the data wrangling process, they trained "Discriminative multinomial naive Bayes" (DMNB) model along with stemming, N-grams tokenizer, and TF-IDF technique. The analyses are carried on a public Twitter data set using a series of performance assessment measures to validate the recommended approach to sentiment analysis. A 10-fold cross-validation is adopted to evaluate the experimental results. Data set used in this study consists of 2000 Arabic tweets classified into positive and negative classes. This study used the WEKA machine learning tool for the experimental setup. The paper also discusses several other machine learning classifiers used in similar work on the same data set with the proposed method's DMNB classifier. Analysis results showed the efficiency of the proposed approach. After the comparison of results the findings shows that proposed model outperformed the models discussed in literature review by improving accuracy of 0.3%.

Duwairi et al. [77] discuss the supervised learning model for sentiment analysis of Arabic text. A data set was compiled and labeled using crowd sourcing, consisting of 2591 Twitter messages. To detect the polarity of a review, SVM, the NB, and KNN classifiers are considered. The data is split into training and validation sets using ten-fold cross-validation. All tweets were pre-processed by removing stop words and removing other unnecessary words. After making tokens of tweets, different weighting schemes were used like TF, TF-IDF, and Bag of Words (BoW). The best precision was 69.97 when the NB classifier used TF. The results of three supervised classifiers were different when different weighting schemes were used. In this study, the K nearest neighbor classifier outperforms other classifiers used with the BoW weighting scheme. When TFIDF was used, SVM obtained the best results. Finally, when TF was used, NB gave the best performance.

Ismail et al. [78] proposed a framework using supervised learning on sentiment analysis for the Arabic language. The Sudanese Arabic dialect corpus was used in this study that was consisted on 4712 tweets. These tweets were compiled and manually labeled by 3 native Arabic individuals for improving labeling process. KNN, SVM, NB, and multi-nominal logistics regression models were trained and tested on the collected

data. The KNN classifier outperforms other supervised learning models by securing an accuracy of 92% when the value of $K = 2$.

Alsaleem [79] presents an article on sentiment analysis using SVM and NB. They use two supervised learning algorithms named SVM and NB. After the model building, both models were compared based on precision-recall and accuracy. The study's findings that the SVM classifier outperforms the NB classifier inaccuracy on the validation data set.

Salamah et al. [80] built a system called Kuwaiti Dialect Opinion Extraction System from Twitter (KDOEST). They use a decision tree and SVM classifier for their data set. By dividing the Kuwaiti words into classes such as happiness class, they extracted features for the system. Such groups are listed as positive or negative. This study shows that SVM performs better than the decision tree for the classification of Arabic tweets.

Al-Osaimi [81] used a supervised approach, where Rapid Miner was used to classifying their gathered tweets using both NB and Decision tree algorithms. They investigate the impact of considering the emotional faces (emoji) typically used by users of Twitter. Their technique showed that the classification of emotion faces increased models' accuracy from 58.28% to 63.79%.

Abdul-Mageed et al. [82] present research on sentiment analysis of Arabic tweets. They used SVM supervised machine learning model for classification. A total of 3015 Arabic tweets was used as a data set in this paper extracted from the TAGREED corpus.

Amira et al. [83] present a study on sentence-level sentiment analysis for Arabic text. Using 2000 tweets from Twitter, they investigated the supervised machine learning techniques for the classification of sentences. In their work, they use a SVM and NB approach. Their work indicate that SVM outperforms the NB classifier.

Oussous et al. [84] present a novel framework for Arabic sentiment analysis. For both Arabic text pre-processing and Arabic sentiment analysis, their system incorporates different approaches and productive models. A novel Arabic data set "Moroccan Arabic data" was created during this study that contain 2000 tweets. This data set contains balanced ratio of positive and negative opinion. Initially the data set contains Informal structures, non-standard dialects, and several other unnecessary words that were removed in the later stages of the study. This data was gathered from various users and real resources. The feature vectors of tweets were pre-processed in many ways and the impact of these on the accuracy of the classifiers have been investigated in this study. The findings indicate that elimination, normalization, and stemming of stop words marginally improved the classification's accuracy. Furthermore, the experimental results showed that in the case of Arabic sentiment analysis, deep learning models, that is, CNN and LSTM are more effective and have shown better performance than machine learning models like SVM, NB. In all these scenarios, deep learning models worked better than conventional models: using unigrams, using stop words, without stop words, with stemmed or without stemmed words.

Ombabi et al. [85] present a deep learning-based approach for Arabic sentiment analysis using social media data. The multi-domain corpus was used in this study. The results showed the impressive performance of their model with 89.10%, 92.14%, 92.44% and 90.75%, accuracy respectively on multi domain corpus. The effect of word embedding approaches on the characterization of Arab sentiment has been thoroughly validated by this review. The findings of this research also indicate that the Fast/Text model is a significant alternative to semantic and syntactic data learning.

Discussion and Learned Lessons

Supervised machine learning techniques have been widely used for Arabic text classification. Different researchers use various classifiers for Arabic text classification as a comparison is shown in Table 2. Most of the studies use supervised learning techniques (i.e., NB and SVM) for text classification on various data sets. These two classifiers outperform other classifiers on the basis of accuracy, precision, and recall. SVM is a supervised machine learning algorithm that uses support vectors and hyperplanes to classify objects

into different classes. Duwari et al. [71], Atoum et al. [72], Duwari [77] and many other researchers used SVM for the Arabic tweets classification on various amount of tweets. They show that SVM perfectly classifies the Arabic text by gaining a handsome level of accuracy. Ismail et al. [78] got the highest accuracy of 92% using SVM classifier on Sudanese Arabic dialect corpus. The second classifier that outperforms other classifiers for Arabic text classification is NB. Naive Bayes classifiers work based on Bayes theorem. For the classification of objects NB classifiers assume strong independence between various attributes of data points. This classifier is widely used in text classification and medical diagnoses. Alsanad et al. [76], Alsaleem et al. [79], Al-Osaimi et al. [81] used NB classifier along with other classifiers for Arabic text classification on various data sets. From these Al-Osaimi et al. [81], gained a valuable accuracy from other technique. Harrag et al. [86], Motaz et al. [87], Rasheed et al. [88] used the decision tree model and got handsome results on the various corpus of the Arabic language. Another study was proposed by Hammad et al. [89] that used the data obtained from 2000 Arabic reviews from social media for evaluation. They used various supervised machine learning models, including SVM, NB, Feed forward Neural Networks with Back Propagation error (BPNN), and Decision Tree. By using the SVM model, the highest accuracy of 96.06%. To improve the preprocessing phase's accuracy, larger and more complex data sets should also be considered. Abdullah et al. [90] also used NB and SVM models and got an accuracy of 80.3%. We conclude that various supervised machine learning algorithms SVM and NB performs better than other classifiers.

Table 2. Comparison of supervised techniques used for Arabic text classification.

Author	Technique Used	Data Set	Accuracy	Strengths	Weaknesses
Duwari et al. [71]	NB, KNN, SVM	Twitter	75.25%.	10 fold cross validation is used with multiple Classifiers	The data set was very small, and accuracy was low.
Atoum et al. [72]	SVM, NB	Twitter	82.1%	Balanced and unbalance data set is used for all models	Accuracy is not good without light stemming.
Al-Horaibi et al. [74]	NB, Decision Tree	Twitter	64.84%	Data set was annotated by two native Arabic annotators	The Size of tweets were small
Abdelaal et al. [75]	ensemble of surface features	Twitter	88.6%	The bagging, boosting and stacking are used for improving accuracy	It takes more time for training and single data set is used for all algorithms.
Jardaneh et al. [73]	Ada Boost, RF, DT	Twitter	76%	Both content-based and user-based features used	Accuracy was low.
Alasand et al. [76]	(DMNB)	Twitter	88.67%	10-fold cross-validation approach is used	Data set was small
Duwairi et al. [77]	NB, SVM, and KNN	Twitter	69.97%	Weighting schemes like TF, TF-IDF, and BoW were used	Accuracy of the system was not good
Ismail et al. [78]	NB, SVM, and KNN	Sudanese Arabic dialect corpus	92%	Tweets were manually labeled by 3 Arabic speakers.	Huge amount of time is required for manual labeling
Alsaleem [79]	NB, SVM	SNP Arabic	77.9%	Different Arabic data sets were used for training	Classification accuracy was not too good
Salamah et al. [80]	Decision Tree and SVM	Kuaweti language	76%	Manual annotation of Data set	The data set was too big so it increase training time.
Al-Osaimi et al. [81]	NB and Decision tree	Twitter	63.79%.	Detect emotion icons in the tweets	The accuracy of system was low.

Table 2. Cont.

Author	Technique Used	Data Set	Accuracy	Strengths	Weaknesses
Abdul-Mageed et al. [82]	SVM	3015 tweets	69%	Testing with TAGREED corpus	Accuracy was not good
Amira et al. [83]	SVM and NB	Twitter	NA	Uni-grams and combination of Uni-grams and Bi-grams were used	Neutral category was not considered data set very small
Harrag et al. [86]	Decision Tree	Hadith	93%	Two different data sets were used	Low Accuracy on Hadith corpus
Helmy et al. [91]	SVM, BPM	Hadith	96%	Models were used without Machine Translation	Accuracy of BPM was low
Motaz et al. [87]	Decision tree	Al-jazeera news	94%	Combination of preprocessing techniques were used	Data set was small
Rasheed et al. [88]	Decision tree, SVM and NB	Arabic YouTube pages	94.5%	Similarity and sentiment words features were used	Model take more computations

4.2. Unsupervised Machine Learning Techniques

Unsupervised Learning is also a kind of machine learning model in which these models do not require the user's input all the time. Instead, it enables the model to independently discover trends and previously unrecognized details of data [92]. It deals primarily with unlabeled data. This technique is the perfect solution for data exploration, cross-selling techniques, consumer segmentation, and image recognition because of its ability to discover similarities and data differences. Figure 7 illustrates the basic process of unsupervised machine learning.

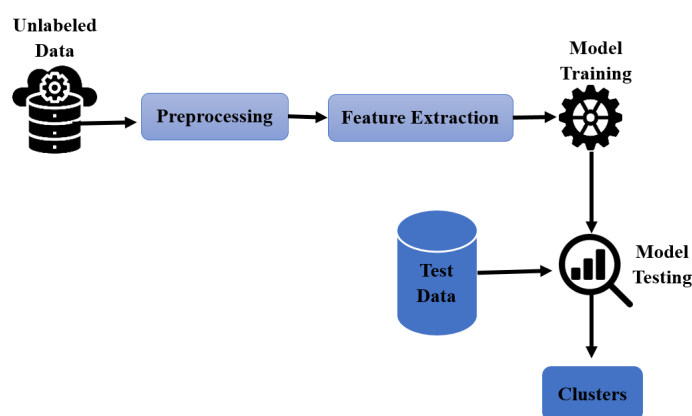


Figure 7. Architecture of Unsupervised Machine Learning.

As compared to supervised learning, using unsupervised learning algorithms, user can process more complicated and complex data [69]. While comparing with other natural methods of learning, it is the more unpredictable model. Unsupervised learning models comprise of anomalies detection, clustering, neural networks. In this section, we discuss various research studies on Arabic text classification.

Al-Azzawy et al. [93] present a novel word clustering strategy for Arabic text. A NLP system was built in this system for the clustering of Arabic words. For this purpose, the K-means clustering technique was used in this study. Finally, they analyze the proposed technique's results with traditional evaluation metrics, that is, Recall, Precision and F-

Measure. The results of the study reveals that the proposed method got the highest accuracy of 98%.

Alzanin et al. [94] focuses on creating efficient Rumor detection from Arabic tweets. They create a training collection using a seed list of phrases that contain rumors. An automatically generated data set was used in this paper. They trained a deep learning classifier based on a character n-gram that can efficiently classify tweets. Results shows that for a small base of labeled data, our semi supervised system E-M outperforms the Gaussian Naive Bayes (NB) with accuracy of 78.6%.

Abuaiadah et al. [95] present an article for clustering Arabic documents by Bisect K-Means technique. This article focuses of measuring efficiency of Bisect K-Means over standard K-means algorithm. Five standard distance measures and 3 stemmers are used in this study. The data set of 300 Arabic documents with nine categories was used for clustering. By using purity measure, the proposed Bisect K-Means outperform the standard K-means algorithm by achieving 92% accuracy. The standard K-means clustering model got the highest accuracy of 88% with Jaccard coefficient function.

Mostafa et al. [96] present a study on Twitter sentiment analysis using unsupervised machine learning. In this study, they analyze a random sample of 3919 halal food tweets. A predefined expert lexicon of 6800 seed adjectives was used. A generally beneficial sentiment towards Halal food was defined by descriptive statistical analysis. Simultaneously, clustering partitioning around medoids (PAM) suggested that halal food consumers can be clustered into four separate segments. They found that halal food consumers constitute an extremely heterogeneous group, divisible by the degree of self-identity, religiosity, animal welfare attitudes, and importance for food quality.

Sangaiah et al. [97] propose an unsupervised technique for clustering Arabic text. After the pre-processing of text, they use incremental k means, k means, and K means with dimensionality reduction for the Arabic text's clustering. They then apply the term weighting method to obtain every term's weight concerning its text. F-measure and entropy are used for calculating accuracy in this study. The accuracy of proposed methods compared to other techniques and the methods proposed indicates greater accuracy and fewer errors in the current classification test cases. Considering that dimension reduction is very sensitive, increasing the reduction ratio will damage essential factors.

Abuaiadah et al. [98] present an article on Twitter sentiment analysis using a clustering approach. In this study, short Arabic sentences are clustered into mainly two categories, that is, positive and negative, for sentiment analysis. This study focuses on clustering Arabic tweets by using linguistic preprocessing and similarity functions. The K-Means clustering algorithm was used for clustering Arabic tweets into two categories. After the analysis of results, it was found that root based stemming is far better than light stemming. The similarity function "Average Kullback-Leibler Divergence" outperforms Pearson Correlation, Cosine, Jaccard Coefficient and Euclidean functions. The results show that Average Kullback-Leibler Divergence with root based stemming got 76% accuracy.

A article was presented by Elarnaoty [99] on Arabic opinion holder extraction. This paper presents a leading analysis independent of any lexical parser for the opinion holder extraction in Arabic news. They researched the creation of a comprehensive feature set to compensate for the lack of structural results in parsing. The suggested feature collection, their proposed semantic field and named entities features were tuned from English articles. Their proposed work was based on Conditional Random Fields (CRF) and semi-supervised pattern recognition techniques. The results of study reveals that the proposed model got 54.03 F-measure score. Oraby et al. [100] proposed a rule based technique for sentiment analysis of Arabic text. Language-specific traits were used that are valuable to segment a text syntactically. A rule-based methodology for opinion-pharse extraction was implemented using an adapted sentiment lexicon and sets of opinion indicators. A new method was introduced for measuring the opinion strength and parsed opinions in this study.

Discussion and Learned Lessons

Unsupervised machine learning algorithms are also used for Arabic text classification. Table 3 presents the summary of Unsupervised machine learning approaches used for Arabic text classification. Various unsupervised models, that is, K-means clustering, KNN, association rule learning, have been used with various Arabic data sets for text classification in the last decade. From different unsupervised approaches, K-means outperform mostly, K-means is the simplest unsupervised learning algorithm to solve clustering problems. The method follows a simple and efficient way to identify the data set by a number of clusters. The key concept is to identify k centers for each cluster. In our survey Sangaiah et al. [97], Abuaiadah et al. [98] used K-means clustering algorithm for the problem. Elarnaoty et al. [99], and Oraby et al. [100] also used unsupervised techniques for the classification of Arabic news articles and movies reviews. El-Halees [101] also proposed a combined classification technique for Arabic opinion mining using the Twitter data set. Huang et al. [102] present an improved unsupervised technique for Arabic dialect. They used social media data and the results of their study show improvement in the accuracy of discussed models by 5%.

Table 3. Comparison of Unsupervised techniques used for Arabic text classification.

Author	Technique Used	Data Set	Accuracy	Strengths	Weaknesses
Salloum et al. [103]	Unsupervised approach	Custom data set	92%	DA-English parallel corpus's was used	Only morphological segmentation of text is considered
Alzanin et al. [94]	Gaussian Naïve Bayes	Twitter	78.6%	Rumor Detection in Arabic language	The Data set was small.
Abuaidah et al. [95]	K-means,	Twitter	98%	Five similarity functions were used	Data set was small
Mostafa et al. [96]	clustering	Twitter	77%	Predefined expert lexicon of 6800 seed adjectives was used	Sarcastic expressions was not detected by the system
Sangaiah et al. [97]	K means with dimensionality reduction	Custom data set	82%	Term weighting method was used	Increase in the reduction ratio damage essential factors
Alotaibi et al. [104]	K-Nearest Neighbors	News Reviews	82%	Word clustering is used	Computation is increased by using word clustering
Abuaiadah et al. [98]	Standard K-means, K-means algorithms	Twitter	76.4%	Root based stemming is used that requires less memory usage	Data set was small and accuracy was not too good.
Elarnaoty et al. [99]	CRF	Arabic news text	85.52	Sequential tagging is used	Arabic Opinion Holder Extraction task performance can be improved by robust Arabic lexical parser
Oraby et al. [100]	Rule-based	Arabic movie reviews	N/A	Basic decomposition and modeling of the Arabic grammatical structure.	Small data set

4.3. Hybrid Machine Learning Techniques

The hybrid machine learning technique is usually based on integrating two different machine learning techniques [105]. A hybrid classification model can consist of a supervised machine learning model for pre-processing data and an unsupervised learning model to make insights or make clusters of data or vice versa [106]. Figure 8 shows the process of the hybrid machine learning technique.

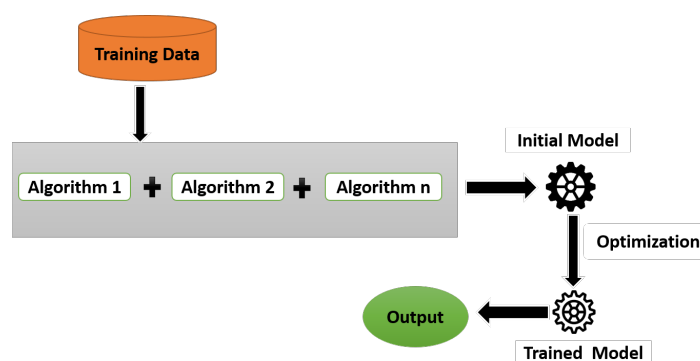


Figure 8. Architecture of Hybrid Machine learning.

Hybrid approaches can also be used for Arabic language processing. Since the last decade, a substantial amount of work has been performed on Arabic language processing using hybrid approaches. Let us explore some of the work done in the Arabic language using hybrid techniques [107]. This section explores some research work done on Arabic text classification.

Aldayal et al. [108] proposed a sentiment analysis technique for the Arabic language by using a hybrid approach. The Twitter data set is used for sentiment analysis in this article. They use a hybrid approach for evaluating the advantages and disadvantages of machine learning approaches and semantic orientation approaches. First, data was collected and passes through the lexicon approach-based classifier. This is used to label the data before giving it to a machine learning classifier for training. SVM is used to classify the tweet into three categories. The proposed hybrid approach got 84.01% classification accuracy. The F-measure of the proposed technique was 84%. The article also discusses the basic features of the Arabic language. They also discuss various challenges of sentiment analysis for the Arabic language.

Thabtah et al. [109] discussed the issue of collecting the emotion from the Arabic text. There are several challenges like the art of tweeting dialectical Arabic, the number of spelling errors, and the vast range of the Twitter domain. The study recommends a hybrid solution that blends semantic orientation and machine learning techniques for sentiment analysis of Arabic tweets. The lexical classification model is used to identify tweets in an unsupervised manner, that is, to deal with unlabeled tweets. The C4.5, PART, RIPPER, one rule classifier are used for Arabic text mining. The study indicates that the least applicable algorithm is OneRule, while the most applicable algorithm in C4.5 beats PART and RIPPER algorithms.

Elshakankery et al. [110] proposed a semi-automatic learning system called HILATSA for sentiment analysis of the Arabic language. HILATSA is a hybrid approach that incorporates techniques focused on both lexicons and machine learning to define the tweets' polarities. This system was capable of handling language changes by using its update property. The HILATSA was divided into three main steps. The first step is called the Lexicon generation step, the second is classification, and the third is called word learner. Initially, different data sets were gathered, and word lexicons were generated from them. Every word of the lexicon is given with three values; PosCount, Neg_Count, and Neu_Count. PosCount describes the frequency of the word considered as positive, the number of times the word deemed to be negative is Neg-Count and NeuCount shows the neutral word count. An emotion lexicon was made based on popular emotions. A separate lexicon was also developed for most common idioms. After this step, machine learning models L2 Logistic Regression, SVM, and Recurrent Neural Network (RNN) classifiers are used to classify sentiments. The accuracy of their system was 83.73% using the recurrent neural network model.

Shaalán et al. [111] propose a hybrid Named Entity Recognition (NER) framework that includes the use of rule-based and machine learning methods to enhance the system's overall efficiency. It tackles the language elicitation bottleneck and the scarcity of tools

requiring deep language processing for underdeveloped languages. They built an Arabic NER framework that can identify 11 forms of Arabic named entities; these named entities include person, place, organization, date, time, price, measurement, percentage, phone number, ISBN, and file name. Empirical findings show that the hybrid solution outperforms all rule-based and ML-based methods as they are separately processed. The proposed hybrid approach got an accuracy of 90% in this article.

Hadni et al. [112] proposed a hybrid approach for Arabic text categorization. In this article, a modern and effective algorithm for Arabic text stemming is proposed to increase the precision of stemming and proposed TC method's accuracy. The proposed framework was a hybridization of three well-known Stemmers. The framework used NB and SVM Classifier for Arabic text categorization to perform text classification. The findings indicate that using the proposed NB stemmer improves the Arabic text categorization efficiency.

AL-Saqqa et al. [113] proposed an ensemble technique for sentiment analysis in the Arabic language. This study focuses on the ensemble voting technique for Arabic text classification. This study used three data sets for the training and evaluation purpose of proposed scheme. These data sets include 500 movie reviews, 2000 tweets and 16,448 book reviews. The combination of NB, SVM, DT, and KNN classifier was used for voting technique using Uni grams and Bigrams. The output of each classifier's is analyzed and compared to the performance of the ensemble voting combination. Multiple experiments have been carried out to test the efficiency of unigram and bigrams. The findings of this study show that the ensemble voting technique outperforms individual models. Besides, the bigram was much more efficient and had better performance than the unigram.

Altaher et al. [114] present a hybrid approach for sentiment analysis of Arabic tweets. They suggest a hybrid approach to the emotion analysis of Arabic tweets focused on deep learning with weighting characteristics. The weighting methods of the features were used in the pre-processing phase to choose the essential elements. Deep learning is an evolving effective strategy to evaluate the feelings of Arabic tweets based on specified features. The experiments indicated the feasibility of the proposed hybrid model, based on deep learning with the chi-square process, and show that the proposed hybrid approach outperforms the SVM, DT, and Neural network classifiers' maximum accuracy and precision efficiency of 90% and 93.7%, respectively.

Biltawi et al. [115] blends lexicon-based and corpus-based methods for Arabic text processing. The goal was to describe the corpus-based system's analysis in the same way as it is done in a lexicon-based method by switching the words of polarity with their respective tags in the lexicon. Three lexicons have been used, including emoji, negation terminology and polarity lexicons. Two tests were carried out with two separate data sets and repeated three times using a common k-fold. Besides, a detailed distinction is made. Experimental results found that their hybrid model surpassed the corpus-based approach, with a maximum precision of 96.34% using random forest with 6-fold cross-validation.

A hybrid approach for sentiment analysis of Arabic text is proposed by Alhumoud et al. [116]. A randomly collected twitter data set is used in this paper for model building. The data set has data of three domains, that is, sports, social, and political. The hybrid learning approach achieved a 6%, 23%, and 21% improvement in accuracy over the classification model SVM in the sports, social and political data sets, respectively. Additionally, with 15% improved accuracy, the KNN classifier in the hybrid learning approach outperformed the supervised model. The results of this study show that the hybrid models outperform supervised models by scoring higher accuracy levels.

Salloum et al. [117] analyze different approaches for Arabic SA-based newspapers and Facebook pages. They consider 24 Gulf newspapers to analyze other text mining and machine learning techniques.

El-Makky et al. [118] present a hybrid approach for classification of Colloquial tweets of the Arabic language. They built a novel annotated data set of the Algerian Egyptian language. This study proposed a newly merged lexicon, an updated semantic orientation

mechanism, and used the information gain measure for feature selection. By using this approach, they improve the accuracy of a hybrid approach by 11%.

Khalifa et al. [119] used a hybrid approach based on lexicon and NB techniques for Arabic text classification. The reviews of Jordan hotels and resorts were used as a data set in this study. After the preprocessing of gathered data, various machine learning models were used with lexicon based model for Arabic question answering. Among all models, the NB model with the lexicon model got the highest macro-average F-measure of 91%. Table 4 present summary of hybrid techniques for Arabic text classification.

Table 4. Comparison of Hybrid techniques used for Arabic text classification.

Author	Technique	Data Set	Accuracy	Strengths	Weaknesses
Aldayal et al. [108]	Semantic orientation + ML	Twitter	84.01%	Basic features of the Arabic language were discussed.	Data set was small
Thabtah et al. [109]	RIPPER + PART	Twitter	83%	The study blends semantic orientation and ML	The training time of the system was high.
Elshakankery et al. [110]	Lexicon + ML technique.	Multiple Data sets	84.6%	Combines both lexical based and machine learning models	The system uses huge amount of computation
Shaanan et al. [111]	Rule based + ML	Twitter	90%	Addresses the bottleneck of language	Data set was small
Hadni et al. [112]	NB + SVM	Kalimat Corpus	94.4%	Three well-known Stemmers were used Tagging	Accuracy of unknown words were low
AL-SAQQA et al. [113]	Four Combinations of SVM, NB and DT	reviews, Tweets	91%	Balanced Data set was used	Neutral category was not considered
Altaher et al. [114]	Deep Learning techniques	Twitter	90%	Deep learning with weighting characteristics were used	The semantic of Arabic tweets was not considered
Biltawi et al. [115]	Dictionary + Corpus based techniques.	Movie Reviews	96.34%	Two separate data sets were used for testing	The lexicon generation takes huge amount of time.
Alhumoud et al. [116]	Custom	Twitter	90%	Data of three Domains, i.e., sports, social, and political were used	Low accuracy on sports data
Salloum et al. [117]	Custom	Facebook, Newspapers	80%	Provides good analysis of Newspapers and Facebook data using ML algorithms.	Data set was very small.
Elzayady et al. [120]	Deep learning models	Education and Politics	80%	Various models were built along with hybrid model	Bidirectional LSTMs can produce good results.
El-Makky et al. [118]	Custom	Twitter	84%	The Arabic lexicon was formed by combining two Modern and two Egyptian Arabic lexicons.	Small data set was used in this study
Khalifa et al. [119]	Lexicon + NB	Reviews on Jordan hotels	91%	Various supervised models were built with lexicon model	Data set was not manually annotated and no standard data set was used

Discussion and Learned Lesson

In the Table 4, we compared different hybrid approaches for Arabic sentiment analysis. In the last decade, various hybrid approaches were proposed for Arabic text classification, especially for Arabic tweet classification. Hybrid techniques outperform other techniques in most cases. However, the time and space complexity of these models is higher than supervised and unsupervised models. Aldayal et al. [108] and Thabtah et al. [109] uses Twitter data set for their hybrid model for tweet classification. Aljarah et al. [105], Al-Smadi et al. [121], Nahar et al. [122], Binsaeed et al. [123] also used hybrid approaches for Arabic language processing using various models and data sets. Elshakankery et al. [110] blend lexicon technique with machine learning techniques on various data set and got a handsome accuracy level of 84%. Shaalan et al. [111], Altaher et al. [114], Alhumoud et al. [116] and El-Makky et al. [118] also used twitter data set on various hybrid techniques for Arabic tweet classification. Hadni et al. [112] used a hybrid approach for their Kalimat Corpus and got 94% accuracy. Biltawi et al. [115], Salloum et al. [117], Khalifa et al. [119] used various Arabic reviews data sets for their hybrid approaches. From all the studies discussed above Biltawi et al. [115] used movies reviews data set and got the highest accuracy level of 96.34%. They used two different data sets for training their hybrid model.

5. Lexicon Based Text Classification

The lexicon approach can be used for text classification of any language. It can be divided into two kinds that is, dictionary based and corpus based techniques [124]. In dictionary based technique first a set of seed words are gathered manually. Then this set of opinion words is extended by using dictionaries. In this step, synonyms and antonym of gathered words is identified and added into the seed list. This process ends when there is no new found in the dictionary. A manual review is also done on the final dictionary to minimize the errors. The Corpus-based techniques access the weaknesses of dictionary based technique. This technique can identify context-specific opinion terms along with seed word list [125]. The identification of such words in the text is done on the basis of syntactic or co-occurrence patterns by using linguistic constraints. In this section, we will explore some work on Arabic text classification using lexicon approaches. Figure 9 illustrates the architecture of a lexicon based technique for text classification.

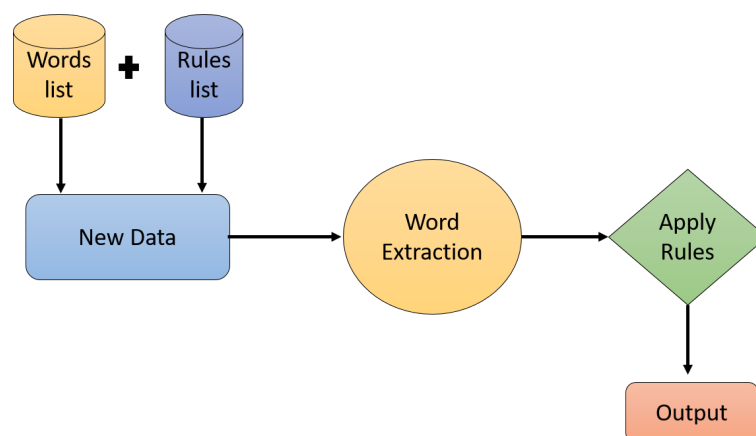


Figure 9. Architecture of Lexicon based Approach.

Al-Ayyoub et al. [126] addressed the text classification problem of the Arabic language by using the lexicon bases technique. The Twitter data set was used in this study and after the preprocessing of tweets, the authors built a sentiment lexicon that contains 120,000 Arabic terms. A sentiment classification tools was built by using that lexicon and predicate calculus. The results of this study show 86.89% accuracy.

Mataoui et al. [127] proposed a lexicon approach for vernacular Algerian Arabic sentiment classification. The proposed method consists of four main modules. These

modules are pre-processing, common phrases similarity computation, language detection & stemming, and polarity computation module. Three different lexicon were built for sentiment classification. These lexicons include negation words, keywords, and intensification words lexicon. Finally, for experimental purposes, they created a test corpus. In order to increase the efficiency of test corpus they manually annotate that corpus. The results of this study show that the proposed model got an accuracy of 79.13% when used with “common phrases similarity computation module”. Common phrases similarity computation, enables the use of common phrases when moving to the term stage. This module compares the text (comment) with the “popular sentences chart” by calculating its parallels (N-gram similarity). If the value of similarity crosses a certain threshold, the module takes the input text into consideration as a usual expression, while no word handling has to proceed to the word.

Duwairi et al. [128] present a novel method for the identification of emotion in Arabic tweets. This approach was based on a sentiment lexicon. The sentiment lexicon was created by translating by Senti-Strength English sentiment lexicon into Arabic. After translation it was extended by using common Arabic phrases. They gathered and manually annotated a collection of 4400 Arabic tweets to test the proposed system’s feasibility. These tweets were categorized into positive and negative tweets using the proposed model according to their sentiment. The proposed model in this study got an accuracy of 70%.

An article was presented by Abdulla et al. [125] for an unsupervised sentiment analysis system of the Arabic language. This article focuses on building a manually annotated data set for Arabic sentiment analysis using corpus bases and lexicon-based approaches. They have done various experiments to test the proposed system’s accuracy and reliability by changing its parameters. After the evaluation, the corpus-based tool outperforms the lexicon-based tool by using SVM to classify a light-stemmed data set of Arabic tweets. Moreover, it was observed that the accuracy of lexicon based technique was improved by increasing the number of lexicons. Badaro et al. [129] proposed a lightweight lexicon based mobile application for Arabic text classification using Twitter data set. They used 3-tier architecture for the classification of tweets into three categories of text that is, Positive, negative, and neutral. A stemmed version of ArSenl was used for the development of the application. User is given with a user interface where he can input text for classification. The proposed approach got an average accuracy of 67.3% on Arabic tweets.

Hmeidi et al. [130] proposed a lexicon based technique for the classification of multi-labeled Arabic text. A total of 8800 documents in a multi label BBC Arabic data set was used in this study. From them 7390 were used for training while 1410 documents were used for testing. The proposed lexicon technique was compared with a corpus based approach by using the MEKA tool. The results of study show that proposed lexicon approach outperforms corpus based technique.

Abdulla et al. [131] proposed a lexicon based Arabic sentiment analysis system. They compare various lexicons and lexicon construction techniques. The proposed model in this study used various novel features like negation words and intensification. The results of the study show that the proposed model got an accuracy of 74.6%. Table 5 presents the comparison of proposed lexicon-based techniques for Arabic text classification.

Table 5. Comparison of Lexicon based techniques used for Arabic text classification.

Author	Technique	Data Set	Accuracy	Strengths	Weaknesses
Alsmadi et al. [132]	Dictionary based	Twitter	86.89%	Predicate calculus was used for text classification	Data set was too big that take long time.
Abdulla et al. [125]	Dictionary and Corpus based	Custom	84.4%	Both techniques were used	Data set was small.
Hamed et al. [127]	Dictionary based	Algerian Arabic Corpus	78.13%	Three different lexicon were used	Accuracy was low.

Table 5. Cont.

Author	Technique	Data Set	Accuracy	Strengths	Weaknesses
Duwairi et al. [128]	Corpus based	Twitter	70%	Senti-Strength English sentiment was used for translating tweets	Low accuracy level.
Abdullah et al. [125]	Clustering	Manual data set	87 %	Corpus and lexicon	No standard data set was used.
Badaro et al. [129]	Corpus based	Twitter	67.3%	Mobile App was built	Low accuracy level
Alayba et al. [133]	Word embeddings	Al-Khair Corpus	92%	Various word2vec Models were built	High computation because of very large data set
Abdulla et al. [131]	Dictionary based	Twitter	74.6%	Various novel features i.e., negation and intensification were used	Accuracy was low
Kabi et al. [134]	Dictionary based	Twitter, Yahoo-Maktoob	70.05%	Data was analysed on Both Light stemming and No stemming	Accuracy on Yahoo corpus were very low i.e 63%
Hossam et al. [135]	Corpus based	Arabic Tweets/ Product reviews/ Hotel Reviews	95%	Various data sets were used	System was not able to detect sarcasm and some idioms
Mohammad et al. [136]	Corpus based	Twitter	66.57%	Arabic text was translated and then classified	Low accuracy score
Al-Ayyoub et al. [126]	Dictionary based	Twitter	86.89%	Large lexicon of words was used	The approach did not handle domain specific issues

Discussion and Learned Lessons

Like machine learning techniques, lexicon based approaches are also widely used for text classification of the Arabic language as compared in Table 5. Most of the proposed studies used dictionary based approaches for Arabic text classification. This survey explores various dictionary and corpus based techniques proposed for Arabic text classification. Various researchers Al-Ayyoub et al. [126], Mataoui et al. [127], Abdulla et al. [131], Abdulla et al. [131], Kabi et al. [134], used dictionary-based approaches on different standard data sets for discussed problem. Al-Smadi et al. [132] achieved the highest accuracy level on a Twitter data set. One reason behind their successful technique is that they used predicate calculus for text classification while no other researcher used it. The authors used both dictionary and corpus based techniques and got handsome accuracy score of 84.4% for their twitter data set. Duwairi et al. [128], Badaro et al. [129], Hossam et al. [135] and Mohammad et al. [136] used corpus based techniques for the proposed problem. S. Hossam et al. [135] used Arabic tweets and reviews data and got the highest accuracy of 95%. Aloqaily et al. [137], Nahar et al. [122], Alhammi et al. [138], Touahri et al. [139] and Abdullah et al. [90] also used lexicon based techniques on various data sets for Arabic text classification. We conclude that lexicon techniques can help get insights into data. The dictionary-based techniques perform better than corpus based approaches for the discussed problem.

6. Challenges of Arabic Text Classification

There are various challenges and limitations for Arabic text classification. These challenges should be properly assessed for making efficient models for Arabic language processing. In this section discusses various challenges for Arabic text classification.

6.1. Small Number of Comprehensive Data Sets

Arabic is a language that is spoken by more than 400 million people around the globe. However, there is very limited data on this language on the internet. There are very few data sets available for Arabic text classification as compared to the English language. That makes it hard to compare output among languages, as the accuracy of text classification highly depends on the amount of data.

6.2. Sarcasm in Text

Sarcasm is a significant issue in text classification. Due to the non-detection of sarcasm accuracy of NLP systems becomes ambiguous. Sarcasm detection is a very difficult issue and requires an intelligent system. There has been limited research on this problem. This challenge should be properly assessed for improving the accuracy of Arabic text classifiers.

6.3. Compound Phrases And Idioms

Compound phrases and idioms are widely used in different Social media platforms such as Twitter and Facebook. Such phrases may vary from one dialect to another and may have a different meaning in various regions of the Arab world. This problem leads to the need for other models for different areas of the Arab world. Also, new phrases are evolving in the Arabic language day by day so it is hard for Arabic text classifiers to classify these phrases accurately.

6.4. Arabizi

Arabizi is a recent social media trend in which a person uses Arabic words to express Latin characters. In addition, many Arabic users prefer to toggle between Arabic and English, finding it challenging to determine whether a word is typed in Arabiz or English. This challenge has also many adverse effects on Arabic classification and it is not widely explored in research.

6.5. Repetition of Words

when there are repeated words in Arabic text, it cannot occur more than twice in Arabic. Thus, if the repetition occurs more than two times at the beginning, middle or end of the word, this will be identified in the pre-processing phase. Unfortunately, if a word is repeated just two times, repetition cannot be identified.

6.6. Negations

Another challenge to Arabic text classification is ignoring negations. Due Arabic negations, word polarities are greatly affected. Informal Arabic has several informal negation words that often influence the polarities of the text by converting the context of the text to completely the opposite. In addition, Arabized expressions are also used, as discussed above in informal Arabic. In informal Arabic, Arabic words are often used as negative words.

6.7. Complex Morphology

There are a variety of dialects in the Arabic language and it has complex morphology. Because of this it needs advanced pre-processing and Lexicon-building techniques. And because of various dialects, the Arabic data that is available online may have words with different meanings.

7. Deep Learning for Arabic Sentiment Analysis

Deep learning has made significant progress in the area of emotion analysis in the English language. However, there has been less study into using deep learning in Arabic sentiment analysis. Recursive Neural Tensor Networks were used in a recent study [140] in the Arabic language, which generated state-of-the-art findings over previous linear models (RNTN). We investigate the model used in [141], which took first place in the SemEval 2017 competition. This model was created for English data and produced cutting-edge outcomes. To forecast the mood of tweets, the algorithm blends CNN and LSTM models. It does not need any extra feature engineering since it uses pre-trained word embeddings. Deep learning (DL) approaches have been evolved in the current decade for Arabic SA. These techniques perform better than traditional ML techniques. Various researchers proposed different DL models for Arabic SA. Some Other DL studies are as follows.

Al Sallab et al. [142] proposed the DL model for Arabic SA. They used three DL architectures, that is, DNN, DBN, and Deep Auto Encoders, for their Linguistic Data Consortium Arabic Tree Bank (LDC ATB) data set. The sentiment scores from the ArSenL lexicon were used for the features vector. They evaluate proposed models by finding accuracy and F1 score. They found that the Deep Autoencoder model provides a more accurate representation of the sparse input vector. They also suggested a fourth algorithm, RAE, which was the best deep learning model based on their data, although it did not require a sentiment lexicon. Boukil et al. [40] also proposed DL based technique for Arabic text classification. They suggest a simple and effective approach for classifying Textual data from large datasets. As a baseline, they evaluate the dataset using CNNs and some traditional ML models. They conclude that the CNNs model outperforms traditional models and achieve a higher level of accuracy.

Mohammed et al. [143] proposed DL models, that is, CNN, RCNN and LSTM models for Arabic SA on 40K Arabic tweets. The data was preprocessed and applied to DL models for text classification. They found that the LSTM model outperforms other models by achieving 88% accuracy. Ombabi et al. [85] also used CNN and LSTM model for Arabic SA on social media data. They used multi-domain corpus and evaluated models on the basis of precision, recall and F1 score. The proposed models were compared with various ML models and found that the proposed model outperforms ML models by achieving 90.75% accuracy level. Omara et al. [144] proposed two deep CNNs for Arabic SA by using only character level features. To train networks, a large scale dataset was built from available SA datasets. The dataset contains opinions from various domains articulated in various Arabic styles (Modern Standard, Dialectal). In addition, various ML such as Logistic Regression, SVM, and NB were used to evaluate the results on a large dataset. The study results indicate that DL models got an enhanced accuracy of 7% compared to ML models.

8. Transformer for Arabic Text

Transformers for NLP explores deep learning for machine translations, speech-to-text, text-to-speech, language modelling, query answering, and many other NLP domains in context with the transformers in great detail. Recently transformers are widely used for Arabic SA and sarcasm detection. Chowdhury et al. [145] present a transformers based technique for Arabic text categorization. They evaluated the impact of pre-training a BERT model on a combination of formal and informal data on text categorization to BERT trained models exclusively on formal text. The main finding in their research is that expanding the training data, whether by using diverse training data for a given task or by using diverse data to pre-train a BERT model, contributes to overall improvements in classification. Farha et al. [146] present an article on transformer-based Language Models to Assess Arabic SA and Sarcasm Detection. They analyze the reliability of 24 models on Arabic SA and sarcasm detection in this article. According to their findings, the models that perform the best are those that have been trained on only Arabic data, including dialectal Arabic, using a greater number of parameters, like the recently introduced MARBERT.

Moreover, they discovered that AraELECTRA is among the best performing models despite much lower computation complexity.

Abuzayed et al. [147] present BERT based technique for Arabic SA and sarcasm detection. This article worked with 7 BERT-Based models and supplemented the shared task data collection to classify tweet's sentiment or recognize sarcasm. With data augmentation, they achieved promising results for sarcasm detection and emotion recognition using the MARBERT model. Abdul-Mageed et al. [148] address the issues of Arabic SA by implementing two efficient deep bidirectional transformer-based architectures, that is, ALBERT and MARBERT. They suggest ArBench, a recent benchmark for multi-dialectal Arabic comprehension, for testing their models. bench was designed with 41 datasets that target five separate tasks clusters, enabling them to have a sequence of structured experiments in rich conditions. According to [148] "When fine-tuned on ArBench, ARBERT and MARBERT collectively achieve new SOTA with sizeable margins compared to all existing models such as mBERT, XLM-R (Base and Large), and AraBERT on 37 out of 45 classification tasks on the 41 datasets".

9. Future Research Directions

Arabic sentiment analysis is a very promising research area, especially because social media data extracting and analysis are becoming very famous.

- By using deep learning, a new hybrid approach can be developed. Big data applications and technology, such as MapReduce and Hadoop, can solve any of the current problems in Arabic sentiment analysis.
- Research and study of sentiments as highlighted in this survey to get the optimal Arabic Sentiment Analysis (ASA) method.
- Most of the techniques rely on manually assembled resources; it is needed to propose new systems to automatically create resources automatically.
- There are several Arabic dialects, mostly these is processed individually. We need to propose methods and techniques that can process all dialects.
- In most of the research studies, researchers follow the way of construction. It is needed to find a way, how to use existing resources for the construction.
- Deep learning approaches are very much promising in different fields of human life, health care, agriculture, image processing. A little work is done on these Arabic text techniques, so it is a very promising area to find appropriate deep learning methods for Arabic text processing.
- For the English language, several applications are operating based on the NLP paradigm, in contrast, Arabic text processing does not get much importance. To fill this gap, the research community should target such applications to process Arabic text.
- Due to the large-scale usage of the internet and social media, a new form of Arabic text is evolved known as Arabizi (derived from Arabic dialects speaking and written in Latin words). This Arabizi is widely used in tweets, it is needed to work on the detection and analysis of these tweets.
- Some enterprise tools and software should be developed for Arabic text to enhance different product sales by analyzing user comments and reviews.
- More large dictionaries and data sets should be considered for Arabic text analysis.
- A big corpus can be built that have multi dialect Arabic data and used for the evaluation purposes of techniques.
- Hybrid models can be used for the detection of negation in text for more reliable results.
- More research should be carried on semantic analysis, because the same word may have multiple meanings in different contexts.

10. Conclusions

Due to the abundant use of Twitter, the retrieval of Arabic dialect is becoming a very complex process. These tweets have valuable information for decision-making to fetch recent trends, especially for government agencies, manufacturing units, social media

observers, and so forth. We classified the state of art research studies according to different machine learning classifications and lexicon-based classifications. This survey reviewed various research studies proposed for Arabic text classification and finds new research areas for the future—different supervised, unsupervised, and lexicon-based techniques for Arabic Tweet classification. Moreover, the limitation and challenges of various approaches in Arabic Tweet classification were identified in this survey. SVM is the most suitable model for Arabic text analysis after analyzing and comparing supervised machine learning techniques. Naive Bayes can also be used for getting a high level of accuracy for the high dimensionality of inputs. The K-means clustering model is widely used for the discussed problem and got handsome results from unsupervised machine learning techniques. Various hybrid approaches for Arabic Tweets were analyzed and it was observed that hybrid models outperform supervised and unsupervised models in most cases.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Greenwood, S.; Perrin, A.; Duggan, M. Social media update 2016. *Pew Res. Cent.* **2016**, *11*, 1–18.
- Asur, S.; Huberman, B.A. Predicting the future with social media. In Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, Toronto, ON, Canada, 31 August–3 September 2010; Volume 1, pp. 492–499.
- Fuchs, C. *Social Media: A Critical Introduction*; Sage: Thousand Oaks, CA, USA, 2017.
- Tartir, S.; Abdul-Nabi, I. Semantic sentiment analysis in Arabic social media. *J. King Saud-Univ.-Comput. Inf. Sci.* **2017**, *29*, 229–233. [\[CrossRef\]](#)
- Hughes, D.J.; Rowe, M.; Batey, M.; Lee, A. A tale of two sites: Twitter vs. Facebook and the personality predictors of social media usage. *Comput. Hum. Behav.* **2012**, *28*, 561–569. [\[CrossRef\]](#)
- Griffis, H.M.; Kilaru, A.S.; Werner, R.M.; Asch, D.A.; Hershey, J.C.; Hill, S.; Ha, Y.P.; Sellers, A.; Mahoney, K.; Merchant, R.M. Use of social media across US hospitals: Descriptive analysis of adoption and utilization. *J. Med. Internet Res.* **2014**, *16*, e264. [\[CrossRef\]](#)
- Elnagar, A.; Al-Debsi, R.; Einea, O. Arabic text classification using deep learning models. *Inf. Process. Manag.* **2020**, *57*, 102121. [\[CrossRef\]](#)
- Abd Al-Aziz, A.M.; Gheith, M.; Eldin, A.S. Lexicon based and multi-criteria decision making (MCDM) approach for detecting emotions from Arabic microblog text. In Proceedings of the 2015 First International Conference on Arabic Computational Linguistics (ACLing), Cairo, Egypt, 17–20 April 2015; pp. 100–105.
- Neri, F.; Aliprandi, C.; Capeci, F.; Cuadros, M.; By, T. Sentiment Analysis on Social Media. *ASONAM* **2012**, *12*, 919–926.
- Yu, Y.; Duan, W.; Cao, Q. The impact of social and conventional media on firm equity value: A sentiment analysis approach. *Decis. Support Syst.* **2013**, *55*, 919–926. [\[CrossRef\]](#)
- Yue, L.; Chen, W.; Li, X.; Zuo, W.; Yin, M. A survey of sentiment analysis in social media. *Knowl. Inf. Syst.* **2019**, *60*, 617–663. [\[CrossRef\]](#)
- Al-Radaideh, Q. Applications of Mining Arabic Text: A Review. In *Recent Trends in Computational Intelligence*; IntechOpen: London, UK, 2020.
- Shehab, M.A.; Badarneh, O.; Al-Ayyoub, M.; Jararweh, Y. A supervised approach for multi-label classification of Arabic news articles. In Proceedings of the 2016 7th International Conference on Computer Science and Information Technology (CSIT), Amman, Jordan, 13–16 July 2016; pp. 1–6.
- Ahmed, N.A.; Shehab, M.A.; Al-Ayyoub, M.; Hmeidi, I. Scalable multi-label arabic text classification. In Proceedings of the 2015 6th International Conference on Information and Communication Systems (ICICS), Amman, Jordan, 7–9 July 2015; pp. 212–217.
- Joulin, A.; Grave, E.; Bojanowski, P.; Mikolov, T. Bag of tricks for efficient text classification. *arXiv* **2016**, arXiv:1607.01759.
- Abdullah, M.; Hadzikadic, M. Sentiment analysis on arabic tweets: Challenges to dissecting the language. In Proceedings of the International Conference on Social Computing and Social Media, Vancouver, BC, Canada, 9–14 July 2017; pp. 191–202.
- Al-Moslmi, T.; Omar, N.; Abdullah, S.; Albared, M. Approaches to cross-domain sentiment analysis: A systematic literature review. *IEEE Access* **2017**, *5*, 16173–16192. [\[CrossRef\]](#)
- Almuqren, L.; Alzammam, A.; Alotaibi, S.; Cristea, A.; Alhumoud, S. A review on corpus annotation for Arabic sentiment analysis. In Proceedings of the International Conference on Social Computing and Social Media, Vancouver, BC, Canada, 9–14 July 2017; pp. 215–225.
- Alnawas, A.; Arici, N. The corpus based approach to sentiment analysis in modern standard Arabic and Arabic dialects: A literature review. *Politek. Derg.* **2018**, *21*, 461–470. [\[CrossRef\]](#)
- Alhumoud, S.O.; Altuwaijri, M.I.; Albuhairi, T.M.; Alohaideb, W.M. Survey on arabic sentiment analysis in twitter. *Int. Sci. Index* **2015**, *9*, 364–368.

21. Assiri, A.; Emam, A.; Aldossari, H. Arabic Sentiment Analysis: A Survey. *Int. J. Adv. Comput. Sci. Appl.* **2015**, *6*. [\[CrossRef\]](#)
22. Al-Ayyoub, M.; Nuseir, A.; Alsmearat, K.; Jararweh, Y.; Gupta, B. Deep learning for Arabic NLP: A survey. *J. Comput. Sci.* **2018**, *26*, 522–531. [\[CrossRef\]](#)
23. Guellil, I.; Saâdane, H.; Azouaou, F.; Gueni, B.; Nouvel, D. Arabic natural language processing: An overview. *J. King Saud-Univ.-Comput. Inf. Sci.* **2019**. [\[CrossRef\]](#)
24. Badaro, G.; Baly, R.; Hajj, H.; El-Hajj, W.; Shaban, K.B.; Habash, N.; Al-Sallab, A.; Hamdi, A. A survey of opinion mining in Arabic: A comprehensive system perspective covering challenges and advances in tools, resources, models, applications, and visualizations. *ACM Trans. Asian-Low-Resour. Lang. Inf. Process. (TALLIP)* **2019**, *18*, 1–52. [\[CrossRef\]](#)
25. Al-Twairish, N.; Al-Khalifa, H.; Al-Salman, A. Subjectivity and sentiment analysis of Arabic: Trends and challenges. In Proceedings of the 2014 IEEE/ACS 11th International Conference on Computer Systems and Applications (AICCSA), Doha, Qatar, 10–13 November 2014; pp. 148–155.
26. Kaseb, G.S.; Ahmed, M.F. Arabic sentiment analysis approaches: An analytical survey. *Int. J. Sci. Eng. Res.* **2016**, *7*, 712–723.
27. El-Masri, M.; Altrabsheh, N.; Mansour, H. Successes and challenges of Arabic sentiment analysis research: A literature review. *Soc. Netw. Anal. Min.* **2017**, *7*, 54. [\[CrossRef\]](#)
28. Dalila, B.; Mohamed, A.; Bendjanna, H. A review of recent aspect extraction techniques for opinion mining systems. In Proceedings of the 2018 2nd International Conference on Natural Language and Speech Processing (ICNLSP), Algiers, Algeria, 25–26 April 2018; pp. 1–6.
29. Hamdi, A.; Shaban, K.; Zainal, A. A Review on Challenging Issues in Arabic Sentiment Analysis. *J. Comput. Sci.* **2016**. [\[CrossRef\]](#)
30. Ghallab, A.; Mohsen, A.; Ali, Y. Arabic Sentiment Analysis: A Systematic Literature Review. *Appl. Comput. Intell. Soft Comput.* **2020**, *2020*. [\[CrossRef\]](#)
31. Abo, M.E.M.; Raj, R.G.; Qazi, A. A Review on Arabic Sentiment Analysis: State-of-the-Art, Taxonomy and Open Research Challenges. *IEEE Access* **2019**, *7*, 162008–162024. [\[CrossRef\]](#)
32. Alsayat, A.; Elmitwally, N. A comprehensive study for Arabic Sentiment Analysis (Challenges and Applications). *Egypt. Inform. J.* **2020**, *21*, 7–12. [\[CrossRef\]](#)
33. Abdul-Mageed, M.; Alhuzali, H.; Elaraby, M. You tweet what you speak: A city-level dataset of arabic dialects. In Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018), Miyazaki, Japan, 7–12 May 2018.
34. Harrat, S.; Meftouh, K.; Smaili, K. Machine translation for Arabic dialects (survey). *Inf. Process. Manag.* **2019**, *56*, 262–273. [\[CrossRef\]](#)
35. Alkhair, M.; Meftouh, K.; Smaili, K.; Othman, N. An arabic corpus of fake news: Collection, analysis and classification. In Proceedings of the International Conference on Arabic Language Processing, Nancy, France, 16–17 October 2019; pp. 292–302.
36. Zeroual, I.; Lakhouaja, A. Arabic corpus linguistics: Major progress, but still a long way to go. In *Intelligent Natural Language Processing: Trends and Applications*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 613–636.
37. Aggarwal, C.C.; Zhai, C. A survey of text classification algorithms. In *Mining Text Data*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 163–222.
38. Ikonomakis, M.; Kotsiantis, S.; Tampakas, V. Text classification using machine learning techniques. *WSEAS Trans. Comput.* **2005**, *4*, 966–974.
39. Kowsari, K.; Jafari Meimandi, K.; Heidarysafa, M.; Mendu, S.; Barnes, L.; Brown, D. Text classification algorithms: A survey. *Information* **2019**, *10*, 150. [\[CrossRef\]](#)
40. Boukil, S.; Biniz, M.; El Adnani, F.; Cherrat, L.; El Moutaouakkil, A.E. Arabic text classification using deep learning technics. *Int. J. Grid Distrib. Comput.* **2018**, *11*, 103–114. [\[CrossRef\]](#)
41. Castillo, C.; Mendoza, M.; Poblete, B. Information Credibility on Twitter. In Proceedings of the 20th International Conference on World Wide Web, WWW '11, Hyderabad, India, 28 March–1 April 2011; Association for Computing Machinery: New York, NY, USA, 2011; pp. 675–684. [\[CrossRef\]](#)
42. Habash, N.; Sadat, F. Arabic preprocessing schemes for statistical machine translation. In *Proceedings of the Human Language Technology Conference of the NAACL, Companion Volume: Short Papers*; Association for Computational Linguistics: New York, NY, USA, 2006; pp. 49–52.
43. Dukes, K.; Habash, N. Morphological Annotation of Quranic Arabic. In Proceedings of the International Conference on Language Resources and Evaluation, LREC 2010, Valletta, Malta, 17–23 May 2010.
44. Traboulsi, H. Arabic named entity extraction: A local grammar-based approach. In Proceedings of the 2009 International Multiconference on Computer Science and Information Technology, Mragowo, Poland, 12–14 October 2009; pp. 139–143.
45. McNeil, K. Tunisian arabic corpus: Creating a written corpus of an ‘unwritten’ language. In *Arabic Corpus Linguistics*; Edinburgh University Press: Edinburgh, UK, 2018; Volume 30.
46. Alansary, S.; Nagi, M.; Adly, N. Building an International Corpus of Arabic (ICA): Progress of compilation stage. In Proceedings of the 7th International Conference on Language Engineering, Cairo, Egypt, 5–6 December 2007; pp. 5–6.
47. Ahmed, M.A.; Hasan, R.A.; Ali, A.H.; Mohammed, M.A. The classification of the modern arabic poetry using machine learning. *Telkomnika* **2019**, *17*, 2667–2674. [\[CrossRef\]](#)
48. Elhassan, R.; Ahmed, M. Arabic text classification on full word. *Int. J. Comput. Sci. Softw. Eng. (IJCSSE)* **2015**, *4*, 114–120.
49. Baier, L.; Jöhren, F.; Seebacher, S. Challenges in the deployment and operation of machine learning in practice. In Proceedings of the 27th European Conference on Information Systems (ECIS), Stockholm & Uppsala, Sweden, 8–14 June 2019.

50. Baltrušaitis, T.; Ahuja, C.; Morency, L.P. Multimodal machine learning: A survey and taxonomy. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *41*, 423–443. [\[CrossRef\]](#)
51. Aggarwal, C.C.; Zhai, C. *Mining Text Data*; Springer Science & Business Media: New York, NY, USA, 2012.
52. Zhang, J.; Zhan, Z.H.; Lin, Y.; Chen, N.; Gong, Y.J.; Zhong, J.H.; Chung, H.S.; Li, Y.; Shi, Y.H. Evolutionary computation meets machine learning: A survey. *IEEE Comput. Intell. Mag.* **2011**, *6*, 68–75. [\[CrossRef\]](#)
53. Pan, W.; Zhong, E.; Yang, Q. Transfer learning for text mining. In *Mining Text Data*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 223–257.
54. Khan, A.; Baharudin, B.; Lee, L.H.; Khan, K. A review of machine learning algorithms for text-documents classification. *J. Adv. Inf. Technol.* **2010**, *1*, 4–20.
55. Das, K.; Behera, R.N. A survey on machine learning: Concept, algorithms and applications. *Int. J. Innov. Res. Comput. Commun. Eng.* **2017**, *5*, 1301–1309.
56. Baydin, A.G.; Pearlmutter, B.A.; Radul, A.A.; Siskind, J.M. Automatic differentiation in machine learning: A survey. *J. Mach. Learn. Res.* **2017**, *18*, 5595–5637.
57. Wang, P.; Li, Y.; Reddy, C.K. Machine learning for survival analysis: A survey. *ACM Comput. Surv. (CSUR)* **2019**, *51*, 1–36. [\[CrossRef\]](#)
58. Benchettara, N.; Kanawati, R.; Rouveiol, C. Supervised machine learning applied to link prediction in bipartite social networks. In Proceedings of the 2010 International Conference on Advances in Social Networks Analysis and Mining, Odense, Denmark, 9–10 August 2010; pp. 326–330.
59. Singh, A.; Thakur, N.; Sharma, A. A review of supervised machine learning algorithms. In Proceedings of the 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), India, New Delhi, 16–18 March 2016; pp. 1310–1315.
60. Cheng, M.Y.; Kusoemo, D.; Gosno, R.A. Text mining-based construction site accident classification using hybrid supervised machine learning. *Autom. Constr.* **2020**, *118*, 103265. [\[CrossRef\]](#)
61. Jaeger, S.; Fulle, S.; Turk, S. Mol2vec: Unsupervised machine learning approach with chemical intuition. *J. Chem. Inf. Model.* **2018**, *58*, 27–35. [\[CrossRef\]](#) [\[PubMed\]](#)
62. Janasik, N.; Honkela, T.; Bruun, H. Text mining in qualitative research: Application of an unsupervised learning method. *Organ. Res. Methods* **2009**, *12*, 436–460. [\[CrossRef\]](#)
63. Goseva-Popstojanova, K.; Tyo, J. Identification of security related bug reports via text mining using supervised and unsupervised classification. In Proceedings of the 2018 IEEE International Conference on Software Quality, Reliability and Security (QRS), Lisbon, Portugal, 16–20 July 2018; pp. 344–355.
64. Huo, H.; Rong, Z.; Kononova, O.; Sun, W.; Botari, T.; He, T.; Tshitoyan, V.; Ceder, G. Semi-supervised machine-learning classification of materials synthesis procedures. *NPJ Comput. Mater.* **2019**, *5*, 1–7. [\[CrossRef\]](#)
65. Wu, C.; Wu, F.; Wu, S.; Yuan, Z.; Liu, J.; Huang, Y. Semi-supervised dimensional sentiment analysis with variational autoencoder. *Knowl.-Based Syst.* **2019**, *165*, 30–39. [\[CrossRef\]](#)
66. Yilmaz, C.M.; Durahim, A.O. SPR2EP: A semi-supervised spam review detection framework. In Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), Barcelona, Spain, 28–31 August 2018; pp. 306–313.
67. Li, Y.; Pan, Q.; Wang, S.; Peng, H.; Yang, T.; Cambria, E. Disentangled variational auto-encoder for semi-supervised learning. *Inf. Sci.* **2019**, *482*, 73–85. [\[CrossRef\]](#)
68. Dalal, M.K.; Zaveri, M.A. Automatic text classification: A technical review. *Int. J. Comput. Appl.* **2011**, *28*, 37–40. [\[CrossRef\]](#)
69. Aggarwal, B.; Mittal, N. Text classification using machine learning methods—a survey. In Proceedings of the Second International Conference on Soft Computing for Problem Solving (SocProS 2012), Jaipur, India, 28–30 December 2012; pp. 701–709.
70. Tong, S.; Koller, D. Support vector machine active learning with applications to text classification. *J. Mach. Learn. Res.* **2001**, *2*, 45–66.
71. Duwairi, R.M.; Qarqaz, I. A framework for Arabic sentiment analysis using supervised classification. *Int. J. Data Min. Model. Manag.* **2016**, *8*, 369–381.
72. Atoum, J.O.; Nouman, M. Sentiment analysis of Arabic jordanian dialect tweets. *Int. J. Adv. Comput. Sci. Appl.* **2019**, *10*, 256–262. [\[CrossRef\]](#)
73. Jardaneh, G.; Abdelhaq, H.; Buzz, M.; Johnson, D. Classifying Arabic tweets based on credibility using content and user features. In Proceedings of the 2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT), Jordan, Amman, 9–11 April 2019; pp. 596–601.
74. Al-Horaibi, L.; Khan, M.B. Sentiment analysis of Arabic tweets using text mining techniques. In Proceedings of the First International Workshop on Pattern Recognition. International Society for Optics and Photonics, Tokyo, Japan, 11–13 May 2016; Volume 10011, p. 100111F.
75. Abdelaal, H.M.; Elmahdy, A.N.; Halawa, A.A.; Youness, H.A. Improve the automatic classification accuracy for Arabic tweets using ensemble methods. *J. Electr. Syst. Inf. Technol.* **2018**, *5*, 363–370. [\[CrossRef\]](#)
76. Alsanad, A. Arabic Topic Detection Using Discriminative Multi nominal Naïve Bayes and Frequency Transforms. In Proceedings of the 2018 International Conference on Signal Processing and Machine Learning, Shanghai, China, 28–30 November 2018; pp. 17–21.

77. Duwairi, R.M.; Qarqaz, I. Arabic sentiment analysis using supervised classification. In Proceedings of the 2014 International Conference on Future Internet of Things and Cloud, Barcelona, Spain, 27–29 August 2014; pp. 579–583.
78. Ismail, R.; Omer, M.; Tabir, M.; Mahadi, N.; Amin, I. Sentiment analysis for arabic dialect using supervised learning. In Proceedings of the 2018 International Conference on Computer, Control, Electrical, and Electronics Engineering (ICCCEEE), Sudan, Khartoum, 12–14 August 2018; pp. 1–6.
79. Alsaleem, S. Automated Arabic Text Categorization Using SVM and NB. *Int. Arab. J. Technol.* **2011**, *2*, 124–128.
80. Salamah, J.B.; Elkhilfi, A. Microblogging opinion mining approach for kuwaiti dialect. In Proceedings of the International Conference on Computing Technology and Information Management (ICCTIM), Dubai, United Arab Emirates, 9–11 April 2014; p. 388.
81. Al-Osaimi, S.; Badruddin, K.M. Role of Emotion icons in Sentiment classification of Arabic Tweets. In Proceedings of the 6th International Conference on Management of Emergent Digital Ecosystems, Buraidah Al Qassim, Saudi Arabia, 15–17 September 2014; pp. 167–171.
82. Abdul-Mageed, M.; Diab, M.; Kübler, S. SAMAR: Subjectivity and sentiment analysis for Arabic social media. *Comput. Speech Lang.* **2014**, *28*, 20–37. [\[CrossRef\]](#)
83. Shoukry, A.; Rafea, A. Sentence-level Arabic sentiment analysis. In Proceedings of the 2012 International Conference on Collaboration Technologies and Systems (CTS), Denver, CO, USA, 21–25 May 2012; pp. 546–550. [\[CrossRef\]](#)
84. Oussous, A.; Benjelloun, F.Z.; Lahcen, A.A.; Belfkih, S. ASA: A framework for Arabic sentiment analysis. *J. Inf. Sci.* **2020**, *46*, 544–559. [\[CrossRef\]](#)
85. Ombabi, A.H.; Ouarda, W.; Alimi, A.M. Deep learning CNN—LSTM framework for Arabic sentiment analysis using textual information shared in social networks. *Soc. Netw. Anal. Min.* **2020**, *10*, 1–13. [\[CrossRef\]](#)
86. Harrag, F.; El-Qawasmeh, E.; Pichappan, P. Improving Arabic text categorization using decision trees. In Proceedings of the 2009 First International Conference on Networked Digital Technologies, Ostrava, Czech Republic, 29–31 July 2009; pp. 110–115.
87. Saad, M.K.; Ashour, W.M. Arabic text classification using decision trees. *Arab. Text Classif. Using Decis. Trees* **2010**, *2*.
88. Elawady, R.M.; Barakat, S.; Nora, M.E. Sentiment analyzer for arabic comments. *Int. J. Inf. Sci. Intell. Syst.* **2014**, *3*, 73–86.
89. Hammad, M.; Al-awadi, M. Sentiment analysis for arabic reviews in social networks using machine learning. In *Information Technology: New Generations*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 131–139.
90. Abdullah, M.; AlMasawa, M.; Makki, I.; Alsolmi, M.; Mahrous, S. Emotions extraction from Arabic tweets. *Int. J. Comput. Appl.* **2020**, *42*, 661–675. [\[CrossRef\]](#)
91. Helmy, T.; Daud, A. Intelligent agent for information extraction from Arabic text without machine translation. In Proceedings of the 1st International Workshop on Cross-Cultural and Cross-Lingual Aspects of the Semantic Web, Shanghai, China, 7 November 2010; Volume 1, p. C3LSW2010.
92. Gentleman, R.; Carey, V.J. Unsupervised machine learning. In *Bioconductor Case Studies*; Springer: Berlin/Heidelberg, Germany, 2008; pp. 137–157.
93. Al-Azzawy, D.S.; Al-Rufaye, F.M.L. Arabic words clustering by using K-means algorithm. In Proceedings of the 2017 Annual Conference on New Trends in Information & Communications Technology Applications (NTICT), Baghdad, Iraq, 7–9 March 2017; pp. 263–267.
94. Alzanin, S.M.; Azmi, A.M. Rumor detection in Arabic tweets using semi-supervised and unsupervised expectation–maximization. *Knowl.-Based Syst.* **2019**, *185*, 104945. [\[CrossRef\]](#)
95. Abuaiaadah, D. Using bisect k-means clustering technique in the analysis of Arabic documents. *ACM Trans. Asian-Low-Resour. Lang. Inf. Process. (TALLIP)* **2016**, *15*, 1–13. [\[CrossRef\]](#)
96. Mostafa, M.M. Clustering halal food consumers: A Twitter sentiment analysis. *Int. J. Mark. Res.* **2019**, *61*, 320–337. [\[CrossRef\]](#)
97. Sangaiah, A.K.; Fakhry, A.E.; Abdel-Basset, M.; El-henawy, I. Arabic text clustering using improved clustering algorithms with dimensionality reduction. *Clust. Comput.* **2019**, *22*, 4535–4549. [\[CrossRef\]](#)
98. Abuaiaadah, D.; Rajendran, D.; Jarrar, M. Clustering Arabic tweets for sentiment analysis. In Proceedings of the 2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA), Hammamet, Tunisia, 30 October–3 November 2017; pp. 449–456.
99. Elarnaoty, M.; AbdelRahman, S.; Fahmy, A. A machine learning approach for opinion holder extraction in Arabic language. *arXiv* **2012**, arXiv:1206.1011.
100. Oraby, S.; El-Sonbaty, Y.; Abou El-Nasr, M. Finding opinion strength using rule-based parsing for arabic sentiment analysis. In Proceedings of the Mexican International Conference on Artificial Intelligence, Mexico City, Mexico, 24–30 November 2013; pp. 509–520.
101. El-Halees, A.M. Arabic opinion mining using combined classification approach. In *Arabic Opinion Mining Using Combined Classification Approach*; Naif Arab University for Security Sciences: Riyadh, Saudi Arabia, 2011.
102. Huang, F. Improved Arabic dialect classification with social media data. In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, Lisbon, Portugal, 17–21 September 2015; pp. 2118–2126.
103. Salloum, S.A.; Al-Emran, M.; Abdallah, S.; Shaalan, K. Analyzing the Arab gulf newspapers using text mining techniques. In Proceedings of the International Conference on Advanced Intelligent Systems and Informatics, Cairo, Egypt, 9–11 September 2017; pp. 396–405.

104. Marie-Sainte, S.L.; Alalyani, N.; Alotaibi, S.; Ghouzali, S.; Abunadi, I. Arabic natural language processing and machine learning-based systems. *IEEE Access* **2018**, *7*, 7011–7020. [\[CrossRef\]](#)
105. Aljarah, I.; Habib, M.; Hijazi, N.; Faris, H.; Qaddoura, R.; Hammo, B.; Abushariah, M.; Alfawareh, M. Intelligent detection of hate speech in Arabic social network: A machine learning approach. *J. Inf. Sci.* **2020**, 0165551520917651. [\[CrossRef\]](#)
106. Silva, E.F.; Barros, F.A.; Prudencio, R.B. A hybrid machine learning approach for information extraction. In Proceedings of the 2006 Sixth International Conference on Hybrid Intelligent Systems (HIS'06), Rio de Janeiro, Brazil, 13–15 December 2006; p. 44.
107. Remeikis, N.; Skucas, I.; Melninkaitė, V. Hybrid machine learning approach for text categorization. *Int. J. Comput. Intell.* **2005**, *1*, 63–67.
108. Aldayel, H.K.; Azmi, A.M. Arabic tweets sentiment analysis—a hybrid scheme. *J. Inf. Sci.* **2016**, *42*, 782–797. [\[CrossRef\]](#)
109. Thabtah, F.; Gharaibeh, O.; Al-Zubaidy, R. Arabic text mining using rule based classification. *J. Inf. Knowl. Manag.* **2012**, *11*, 1250006. [\[CrossRef\]](#)
110. Elshakankery, K.; Ahmed, M.F. HILATSA: A hybrid incremental learning approach for Arabic tweets sentiment analysis. *Egypt. Inform. J.* **2019**, *20*, 163–171. [\[CrossRef\]](#)
111. Shaalan, K.; Oudah, M. A hybrid approach to Arabic named entity recognition. *J. Inf. Sci.* **2014**, *40*, 67–87. [\[CrossRef\]](#)
112. Hadni, M.; Ouatik, S.A.; Lachkar, A. Effective Arabic stemmer based hybrid approach for Arabic text categorization. *Int. J. Data Min. Knowl. Manag. Process.* **2013**, *3*, 1. [\[CrossRef\]](#)
113. Al-Saqqa, S.; Obeid, N.; Awajan, A. Sentiment analysis for Arabic text using ensemble learning. In Proceedings of the 2018 IEEE/ACS 15th International Conference on Computer Systems and Applications (AICCSA), Aqaba, Jordan, 28 October–1 November 2018; pp. 1–7.
114. Altaher, A. Hybrid approach for sentiment analysis of Arabic tweets based on deep learning model and features weighting. *Int. J. Adv. Appl. Sci.* **2017**, *4*, 43–49. [\[CrossRef\]](#)
115. Biltawi, M.; Al-Naymat, G.; Tedmori, S. Arabic sentiment classification: A hybrid approach. In Proceedings of the 2017 International Conference On New Trends In Computing Sciences (ICTCS), Amman, Jordan, 11–13 October 2017; pp. 104–108.
116. Alhumoud, S.; Albuhairi, T.; Altuwaijri, M. Arabic sentiment analysis using WEKA a hybrid learning approach. In Proceedings of the 2015 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K), Lisbon, Portugal, 12–14 November 2015; Volume 1, pp. 402–408.
117. Salloum, S.A.; Mhamdi, C.; Al-Emran, M.; Shaalan, K. Analysis and classification of Arabic newspapers' Facebook pages using text mining techniques. *Int. J. Inf. Technol. Lang. Stud.* **2017**, *1*, 8–17.
118. El-Makky, N.; Nagi, K.; El-Ebshihy, A.; Apady, E.; Hafez, O.; Mostafa, S.; Ibrahim, S. Sentiment analysis of colloquial Arabic tweets. In Proceedings of the ASE BigData/SocialInformatics/PASSAT/BioMedCom 2014 Conference, Harvard University, Cambridge, MA, USA, 14–16 December 2014; pp. 1–9.
119. Khalifa, K.; Omar, N. A hybrid method using lexicon-based approach and Naive Bayes classifier for Arabic opinion question answering. *J. Comput. Sci.* **2014**, *10*, 1961–1968. [\[CrossRef\]](#)
120. Elzayady, H.; Badran, K.M.; Salama, G.I. Arabic Opinion Mining Using Combined CNN-LSTM Models. *Int. J. Intell. Syst. Appl.* **2020**, *4*, 25–36. [\[CrossRef\]](#)
121. Al-Smadi, M.; Al-Zboon, S.; Jararweh, Y.; Juola, P. Transfer Learning for Arabic Named Entity Recognition with Deep Neural Networks. *IEEE Access* **2020**, *8*, 37736–37745. [\[CrossRef\]](#)
122. Nahar, K.M.; Jaradat, A.; Atoum, M.S.; Ibrahim, F. Sentiment analysis and classification of arab jordanian facebook comments for jordanian telecom companies using lexicon-based approach and machine learning. *Jordanian J. Comput. Inf. Technol. (JJCIT)* **2020**, *6*.
123. Binsaeed, K.; Stringhini, G.; Youssef, A.E. Detecting Spam in Twitter Microblogging Services: A Novel Machine Learning Approach based on Domain Popularity. *Int. J. Adv. Comput. Sci. Appl. (IJACSA)* **2020**.
124. Taboada, M.; Brooke, J.; Tofiloski, M.; Voll, K.; Stede, M. Lexicon-based methods for sentiment analysis. *Comput. Linguist.* **2011**, *37*, 267–307. [\[CrossRef\]](#)
125. Abdulla, N.A.; Ahmed, N.A.; Shehab, M.A.; Al-Ayyoub, M. Arabic sentiment analysis: Lexicon-based and corpus-based. In Proceedings of the 2013 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT), Amman, Jordan, 3–5 December 2013; pp. 1–6.
126. Al-Ayyoub, M.; Essa, S.B.; Alsmadi, I. Lexicon-based sentiment analysis of arabic tweets. *Int. J. Soc. Netw. Min.* **2015**, *2*, 101–114. [\[CrossRef\]](#)
127. Mataoui, M.; Zelmati, O.; Boumechache, M. A proposed lexicon-based sentiment analysis approach for the vernacular Algerian Arabic. *Res. Comput. Sci.* **2016**, *110*, 55–70. [\[CrossRef\]](#)
128. Duwairi, R.M.; Ahmed, N.A.; Al-Rifai, S.Y. Detecting sentiment embedded in Arabic social media—a lexicon-based approach. *J. Intell. Fuzzy Syst.* **2015**, *29*, 107–117. [\[CrossRef\]](#)
129. Badaro, G.; Baly, R.; Akel, R.; Fayad, L.; Khairallah, J.; Hajj, H.; Shaban, K.; El-Hajj, W. A light lexicon-based mobile application for sentiment mining of arabic tweets. In Proceedings of the Second Workshop on Arabic Natural Language Processing, Beijing, China, 30 July 2015; pp. 18–25.
130. Hmeidi, I.; Al-Ayyoub, M.; Mahyoub, N.A.; Shehab, M.A. A lexicon based approach for classifying Arabic multi-labeled text. *Int. J. Web Inf. Syst.* **2016**. [\[CrossRef\]](#)

131. Abdulla, N.; Majdalawi, R.; Mohammed, S.; Al-Ayyoub, M.; Al-Kabi, M. Automatic Lexicon Construction for Arabic Sentiment Analysis. In Proceedings of the 2014 International Conference on Future Internet of Things and Cloud, Barcelona, Spain, 27–29 August 2014; pp. 547–552. [\[CrossRef\]](#)
132. Al-Smadi, M.; Talafha, B.; Al-Ayyoub, M.; Jararweh, Y. Using long short-term memory deep neural networks for aspect-based sentiment analysis of Arabic reviews. *Int. J. Mach. Learn. Cybern.* **2019**, *10*, 2163–2175. [\[CrossRef\]](#)
133. Alayba, A.M.; Palade, V.; England, M.; Iqbal, R. Improving sentiment analysis in Arabic using word representation. In Proceedings of the 2018 IEEE 2nd International Workshop on Arabic and Derived Script Analysis and Recognition (ASAR), London, UK, 12–14 March 2018; pp. 13–18.
134. Abdulla, N.A.; Ahmed, N.A.; Shehab, M.A.; Al-Ayyoub, M.; Al-Kabi, M.N.; Al-rifai, S. Towards improving the lexicon-based approach for arabic sentiment analysis. *Int. J. Inf. Technol. Web Eng. (IJITWE)* **2014**, *9*, 55–71. [\[CrossRef\]](#)
135. Ibrahim, H.S.; Abdou, S.M.; Gheith, M. Sentiment analysis for modern standard arabic and colloquial. *arXiv* **2015**, arXiv:1505.03105.
136. Mohammad, S.M.; Salameh, M.; Kiritchenko, S. How translation alters sentiment. *J. Artif. Intell. Res.* **2016**, *55*, 95–130. [\[CrossRef\]](#)
137. Aloqaily, A.; Alhassan, M.; Salah, K.; Elshqeirat, B.; Almashagbah, M. Sentiment analysis for arabic tweets datasets: Lexicon-based and machine learning approaches. *J. Theor. Appl. Inf. Technol.* **2020**, *98*, 114–122.
138. Alhammi, H.A.; Haddar, K. Building a Libyan Dialect Lexicon-Based Sentiment Analysis System Using Semantic Orientation of Adjective-Adverb Combinations. *Int. J. Comput. Theory Eng.* **2020**, *12*. [\[CrossRef\]](#)
139. Touahri, I.; Mazroui, A. Deep analysis of an Arabic sentiment classification system based on lexical resource expansion and custom approaches building. *Int. J. Speech Technol.* **2020**, *24*, 109–126. [\[CrossRef\]](#)
140. Baly, R.; Badaro, G.; El-Khoury, G.; Moukalled, R.; Aoun, R.; Hajj, H.; El-Hajj, W.; Habash, N.; Shaban, K. A characterization study of arabic twitter data with a benchmarking for state-of-the-art opinion mining models. In Proceedings of the Third Arabic Natural Language Processing Workshop, Valencia, Spain, 3 April 2017; pp. 110–118.
141. Cliche, M. BB_twtr at SemEval-2017 task 4: Twitter sentiment analysis with CNNs and LSTMs. *arXiv* **2017**, arXiv:1704.06125.
142. Al Sallab, A.; Hajj, H.; Badaro, G.; Baly, R.; El-Hajj, W.; Shaban, K. Deep learning models for sentiment analysis in Arabic. In Proceedings of the Second Workshop on Arabic Natural Language Processing, Beijing, China, 30 July 2015; pp. 9–17.
143. Mohammed, A.; Kora, R. Deep learning approaches for Arabic sentiment analysis. *Soc. Netw. Anal. Min.* **2019**, *9*, 1–12. [\[CrossRef\]](#)
144. Omara, E.; Mosa, M.; Ismail, N. Deep convolutional network for arabic sentiment analysis. In Proceedings of the 2018 International Japan-Africa Conference on Electronics, Communications and Computations (JAC-ECC), Alexandria, Egypt, 16–18 December 2018; pp. 155–159.
145. Chowdhury, S.A.; Abdelali, A.; Darwish, K.; Soon-Gyo, J.; Salminen, J.; Jansen, B.J. Improving Arabic text categorization using transformer training diversification. In Proceedings of the Fifth Arabic Natural Language Processing Workshop, Barcelona, Spain, 12 December 2020; pp. 226–236.
146. Farha, I.A.; Magdy, W. Benchmarking Transformer-based Language Models for Arabic Sentiment and Sarcasm Detection. In Proceedings of the Sixth Arabic Natural Language Processing Workshop, Kyiv, Ukraine, 19 April 2021; pp. 21–31.
147. Abuzayed, A.; Al-Khalifa, H. Sarcasm and Sentiment Detection In Arabic Tweets Using BERT-based Models and Data Augmentation. In Proceedings of the Sixth Arabic Natural Language Processing Workshop, Kyiv, Ukraine, 19 April 2021; pp. 312–317.
148. Abdul-Mageed, M.; Elmadany, A.; Nagoudi, E.M.B. ARBERT & MARBERT: Deep Bidirectional Transformers for Arabic. *arXiv* **2020**, arXiv:2101.01785.