A Survey of Sentiment Analysis in the Arabic Language

Kamel JAFAR
Master in Web Sciences
Syrian Virtual University
Latakia_Syria
kamel_78564@svuonline.org

Bassel ALKHATIB
Program Director
Web Science_ Artificial Intelligence
Syrian Virtual University_ Damascus University
Damascus_Syria
t_balkhatib@svuonline.org

Hazem WANNOUS
Associate Professor
University de Lille 1 Sciences et Technologies
Telecom Lille
Cité scientifique
59650 Villeneuve-d'Ascq
hazem.wannous@telecom-lille.fr

JDIM

Journal of Digital Information Management

ABSTRACT: *"What the others think?" is an* essential *question for individuals, companies, and governments. All need to know the public opinions to make their decisions wisely. In the last decade, sentiment analysis and opinion mining* have *become one of the growing* research areas*. This paper presents the current state of sentiment analysis and opinion mining research. In particular,* researches those *deals with the Arabic language. We tried to cover the techniques and methods in sentiment analysis and* the challenges *in the field. We described the* leading methods *and approaches that have been introduced in the literature for Arabic Sentiment analysis and opinion mining. The main contributions of this paper include the sophisticated categorizations of a large number of recent articles about Arabic Sentiment analysis. These articles are categorized according to their contributions in the various sentiment analysis techniques.*

Subject Categories and Descriptors: [H.3.1 Content Analysis and Indexing]: Linguistic processing; [I.2.7 Natural Language Processing]: Text analysis

General Terms: Sentiment Analysis, Arabic Language Processing, Sentiment Features

Keywords: Sentiment Analysis, Opinion Mining, Arabic Sentiment Analysis, Sentiment Classification, Feature Selection

Received: 24 August 2021, Revised 9 December 2021, Accepted 23 December 2021

Review Metrics: Review Scale: 0-6; Review Score: 4.52; Inter-reviewer consistency: 81.45%

DOI: 10.6025/jdim/2022/20/1/10-24

1. Introduction

With the emergence of Web 2.0, social networking sites have taken a large part of the everyday life of people around the world, which led them to create a large mass of important information that is useful for individuals, businesses, and governments because everyone needs to know the opinion of others about anything. Individuals want to know which product to buy, so they read reviews to help make their decisions. Companies want to understand how customers feel about their products. Governments want to know public opinions about their existing or proposed policies.

According to the Financial Times report [1], the most critical factor that drove Donald Trump to win in the presidential election was the use of the Internet, social networks, and information about voters to attract them. Information about the voters, compiled by an IT company, includes their psychological characteristics and political tendencies that contributed to dealing with voters properly. The company's analysts said that they had detailed information on the identities of each of the 220 million Americans, collected through the Internet, primarily through social networks. Therefore, one can recognize the importance of sentiment analysis.

Sentiment analysis is a method of capturing people's sentiment (feeling or opinion) towards a specific topic [2]. This field may be considered part of natural language processing, machine learning, and computational linguistics. In other words, it usually tries to evaluate and extract people's sentiment from their writing. SA has many names in literature, including subjectivity analysis, opinion mining, review mining, and appraisal extraction [2]. Recently, sentiment analysis applications have spread to almost every possible domain, from consumer products, health care, tourism, hospitality, and financial services to social events and political elections. [3]

The main goal of sentiment analysis is to recognize the sentiments toward people's topics. Therefore, Sentiment analysis is commonly seen as a sub-area of Natural Language Processing that led to the immense challenge of natural language understanding by machines, especially when we talk about the Arabic language with its morphological complexities and dialectal varieties that require advanced pre-processing and lexicon-building steps. [4]

This paper discusses sentiment analysis problems and related work in general and especially for Arabic sentiment analysis articles. In addition, we provide sophisticated categorization of many recent articles that deal with Arabic sentiment according to the used techniques. In section 2, we introduce the SA problem. Section 3 overviews the SA process and related techniques. Section 4 includes SA Literature Review in general. Arabic SA and its corresponding articles are surveyed in Section 5, and we focused on concept-level and deep learning approaches in sections 6, 7 respectively. Section 8 introduces, in brief, a multi-modal sentiment analysis. In section 9, the results and discussions, and finally, the conclusion and future works are presented in section 10.

2. Sentiment Analysis Problem

In general, any text can be categorized into two types: Fact or Opinion. Facts are objective expressions about entities. Opinions are subjective expressions that describe people's sentiments or feelings toward entities. Opinion on an entity is a positive or negative view from an opinion holder who is the person or organization that expresses the opinion. Opinions can be expressed in the object, which is an entity that can be a product, person, event, organization, or topic [5]. The thing may be having a set of attributes or aspects. The item or its elements (features) are called opinion targets because the opinion can be expressed on the entity in general (called general opinion) or on the specific feature (called specific opinion). Opinion orientation is also known as sentiment orientation, the polarity of opinion, or semantic orientation. It can be positive, negative, or neutral (no opinion). [6]

Sentiment Analysis can be considered as a classification process, in classic topic-based classification documents classified into predefined topic classes, so topic-related words are needed. Still, in sentiment classification, we need opinion words that indicate positive or negative opinions, for example, great, excellent, unique, horrible, bad, worst, etc.

There are three levels in sentiment analysis: [5]

(1) Document-level classifies an opinion document into positive or negative. Researchers assume the document is opinioned and talks about product reviews. In [7], the document's Pre-processing is explored on the dataset. Secondly, Naive Bayes and Support Vector Machines (SVMs) are applied with TF-IDF and BTO (Binary-Term Occurrence). Thirdly, the proposed model for sentiment analysis is expanded.

(2) Sentence level determines whether each sentence expresses a positive, negative, or neutral opinion. At first, in this level, we have to know if the sentence is opinioned or not (Subjective Classification). We can determine if the sentence expresses a positive or negative opinion. Subjective classification classifies sentences into subjective and objective classes [8]. An actual sentence presents factual information about the world, while an emotional sentence expresses personal feelings or beliefs. A subjective sentence can express no opinion; researchers assumed that the subjective sentence is an opinionated sentence and the objective sentence is not a little sentence. However, an objective sentence can express an opinion. An opinionated Sentence is a sentence that expresses explicit or implicit positive or negative opinions.

In the study [9], the researchers attempt to analyze the sentiment at the sentence level first and then use results to analyze the sentiment at the document level. They compared two different approaches. The first was generalizing the Arabic sentence into a general structure that contains the actor and the action. The second used semantic and stylistic features. They used the SVM for the grammatical classifier and obtained an accuracy of 89%.

In comparison, the decision tree was used with the semantic approaches. It achieved an accuracy of 80% when the semantic orientation of the words extracted and assigned manually was used and 62% when the dictionary was used. The process in the study [10] started by getting the tweets from Twitter. Then, they passed by each tweet and labeled it positive or negative. After that, the features in each tweet are extracted and represented in a feature vector. Then, they applied the feature vectors to the NB and SVM Classifiers to compare the results and choose the classifier with the highest accuracy.

(3) Aspect level, many applications for end-user need more fine-grained analysis to know the opinion target for each opinion in the sentence. It is also called feature-based opinion mining.

The work's main contribution in [11] is that three different configurable N-GRAM methods for feature polarity identification are proposed. These methods can be configured with other parameters to obtain the best polarity identification approach.

In the study [12], the researchers created a domain ontology to take advantage of a domain ontology for providing more fine-grained sentiment analysis of Twitter posts regarding the distinct topics of a specific subject discussed in each tweet.

3. Sentiment Analysis Process

Sentiment analysis generally consists of three main steps: preprocessing, feature selection, and sentiment classification.

3.1. Preprocessing

Opinions are available all over the Internet in blogs, forums, and social media websites. Still, the challenge for us as researchers is to aggregate this information and present it for the next step. Preprocessing is an important step, especially for social media text, because it has a lot of misspellings, duplicate characters, and special characters. Pre-processing is usually based on NLP techniques such as tokenization (splitting the sentences into words), de-noising (remove special characters, capture symbols for emotions), normalization (remove duplicate characters, identify root words, etc.), stop-words removal (clear the stop words and the words which are of no use to sentiment analysis), stemming (return the word to its stem or root), lemmatization (convert inflected words to their root form) [13].

The authors of [14] and [15] investigate the role of text preprocessing in sentiment analysis.

The paper of [16] investigates the relevance of using the roots of words as input features into a sentiment analysis system under two distinct domains (Newswire, Movie Reviews) for Modern Standard Arabic and Arabic dialects. They observed that word roots could enhance the sentiment analysis task results on a more generic level instead of using a domain-specific approach.

The paper [17] demonstrated the effectiveness of preprocessing on enhancing the sentiment classification of 1000 Arabic tweets (positive or negative) written from Twitter. In the Machine Learning approach, their stemmer improved 1% over the light stemmer. In the Semantic Orientation approach, their stemmer improved 0.5%.

3.2. Feature selection

The outputs of pre-processing are the extracted text features. Feature selection is considered the bottleneck of the classification process because good features lead to good results. Sound features are informative, independent, and simple.

The main goal of the feature selection is to decrease the dimensionality of the feature space and thus computational cost. As a second objective, feature selection will reduce the over-fitting of the learning scheme to the training data. [18]

Some of the example features are as follows: [5]

• **Terms and their frequency:** Term frequency defines the relative frequency in the document.

• **Binary term occurrence or presence:** term occurrence is defined as the binary value. Binary Term Occurrence = 1 if the term appears at least once in a document, and 0 otherwise. In [19], the term presence shows improvement compared with the term frequency. [20] shows that using binary term instead of term frequency feature does not impact performance.

• **TF-IDF:** it describes how important a word is for a document. It consists of term frequency (TF) and inverted document frequency (IDF).

• **N-grams:** Grams are words that are frequently repeated in the corpus. These features are individual words (unigram), two consecutive words (bigrams), and three successive words (trigrams).

• **Parts of speech (POS):** Like finding adjectives, as they are essential indicators of opinions. Using (POS) tagging system decreases the ambiguity of the word. [21]

• **Sentiment Shifters:** These are expressions used to change sentiment orientations. Negation words are the most critical class of sentiment shifters.

In [20], the authors explored various feature definitions and selection strategies on three datasets of multiple sizes. They show that the size of the dataset affects the performance of some of the techniques, and they present the cost analysis in terms of space used to store the dataset and the time it takes to compute it.

Feature Selection methods: Can be divided into lexicon-based methods and statistical methods.

Lexicon-based approaches usually begin with a small set of words (seed). Then they bootstrap this set through synonym detection to obtain a more extensive lexicon. Statistical techniques are fully automatic. The most frequently used methods are information gain, Gini index, Point-wise Mutual Information (PMI), Chi-square and Latent Semantic Indexing. [22]

In [23], two new feature selection methods are proposed: the first is based on the probability of belongings of a term for a particular class, namely Probability Proportion Difference (PPD). The second one, the Categorical Probability Proportion Difference (CPPD) feature selection method, is proposed based on Probability Proportion Difference (PPD) and Categorical Proportion Difference (CPD). The performance of the proposed feature selection methods is compared with the CPD method and Information Gain (IG). Experimental results show that the proposed methods improve the classification performance from the baseline results very efficiently and filter irrelevant features.

In [24], different feature selection methods are applied for sentiment analysis. They observed Minimum Redundancy Maximum Relevance (mRMR) performs better than Information gain (IG). They compared IG and mRMR for sentiment analysis with a hybrid feature section based on Rough Set Theory (RST) and IG and found that the hybrid approach performs better than IG. Experimental results show that the Hybrid feature selection method with fewer features produces better results than other feature selection methods.

3.2. Sentiment Classification Techniques

Sentiment Classification techniques can be roughly divided into machine learning, lexicon-based, and hybrid approaches.

- **The Machine Learning Approach (ML):** is typically a supervised learning approach in which a set of data labeled (training data) is used to train the ML classifier. Then we can use this classifier to predict the class of unseen or new data called testing data.

There are many kinds of supervised classifiers in literature, can be divided into:

1) Probabilistic classifiers: Naïve Bayes Classifier (NB), Bayesian Network, and Maximum Entropy Classifier (ME).

2) Linear classifiers: Support Vector Machines Classifiers (SVM) and Neural Network (NN).

3) Also, there are Decision Tree Classifiers and Rule-based Classifiers.

- The paper [25] deals with sentiment analysis in Arabic reviews from a machine learning perspective. The Naïve Bayes, SVM, and K-Nearest Neighbor classifiers were run on a dataset of tweets and comments. The results show that SVM gives the highest precision while KNN (K=10) provides the highest Recall.

- For the classification process, the authors of the study [26] used RapidMiner to examine three algorithms: SVM, NB, and KNN. Their work has been done on Jordanian dialect.

- A system for Arabic subjectivity and sentiment analysis was proposed in the paper [27], using a machine learning supervised approach. They used SVM algorithms. They collected Arabic tweets from TAGREED corpus.

- This paper [28] developed a corpus from Arabic Facebook news comments. Then, they applied three different classifiers (Support Vector Machines, Naïve Bayes, and Decision Tree) with other groups of features. The support vector machine classifier obtained the best result.

- In work [29], the authors proposed a new lexicon based on slang sentiment words and idioms, collected manually from news websites, Facebook, and Twitter pages. After that, they applied an SVM-based classifier. The results showed improvement when they used the new lexicon rather than a classical opinion words lexicon.

- For the classification process, the authors of the study [26] used RapidMiner to examine three algorithms: SVM, NB, and KNN. Their work has been done on Jordanian dialect.

- A system for Arabic subjectivity and sentiment analysis was proposed in the paper [27], using a machine learning supervised approach. They used SVM algorithms. They collected Arabic tweets from TAGREED corpus.

- This paper [28] developed a corpus from Arabic Facebook news comments. Then, they applied three different classifiers (Support Vector Machines, Naïve Bayes, and Decision Tree) with other groups of features. The support vector machine classifier obtained the best result.

- In work [29], the authors proposed a new lexicon based on slang sentiment words and idioms, collected manually from news websites, Facebook, and Twitter pages. After that, they applied an SVM-based classifier. The results showed improvement when they used the new lexicon rather than a classical opinion words lexicon.

- **The Lexicon-based Approach:** is an unsupervised learning approach that relies on a sentiment lexicon, a collection of known and precompiled sentiment terms. It is divided into the dictionary-based approach and corpus-based approach, which use statistical or semantic methods to find sentiment polarity to unknown words and determine the polarity of blocks of text. We can further divide unsupervised methods into dictionary-based and corpus-based relative to how the lexicon is built.

- The Dictionary-based approach depends on finding opinion seed words and then searching the dictionary of their synonyms and antonyms. The corpus-based process guarantees context specificity of word orientations by searching a large corpus.

The study [30] aims to compare the accuracy of the two approaches for SA (corpus-based and dictionary-based) and show how the accuracy of the lexicon-based tool improves with the addition of more words to the lexicon. The authors discussed building a manually annotated dataset and then detailed steps of building the lexicon.

The authors in the study [31] proposed a lexicon-based tool. The tool works on Arabic textual opinions, whether they use MSA or colloquial Arabic or both. Therefore, they built manually two general-purpose lexicons to discern the polarity of an opinion expression, whether the opinion uses MSA or/and colloquial Arabic. Another sixteen domain-specific lexicons were built. So the total number of lexicons built is 18, where nine of these polarity lexicons are dedicated to positive polarity, and the other nine lexicons are dedicated to negative polarity. An opinion is considered neutral when its tokens are divided equally between positive and negative lexicons. The tool is capable of determining whether Arabic social media reviews are (subjective or objective), (positive or negative), and (strong or weak).

• **The hybrid Approach:** combines both approaches and is very common with sentiment lexicons playing a vital role in most methods.

The proposed approach in [32] is a hybrid approach that combines a lexicon-based approach with a machine learning approach. The lexicon-based features are used as a part of the feature vector input to the classification module. They worked on colloquial Arabic instead of Modern Standard Arabic, and the conducted performance evaluation experiments showed the outperformance of the system compared to related work.

A hybrid approach proposed in [33] aims to identify the sentiment polarity for collected comments or posts from Twitter. Two machine learning classification techniques are used. They extended the evaluation of prediction algorithms and enhanced them using Bagging and Boosting algorithms. They observed some unique expressions that harm the accuracy of the automatic sentiment prediction system.

The authors of the study [34] proposed a hybrid method of lexicon-based Approach and classification using the Naïve Bayes classifier to address the difficulties of Arabic Opinion Question Answering.

The study in the paper [35] aims to classify Arabic Facebook posts. They applied two approaches, depending on syntactic features, using standard patterns used in different Arabic dialects to express opinions. These patterns achieved high accuracy in determining the polarity of sentiment. The second Approach is an ordinary probabilistic model, a Na G ve-Bayes classifier that assumes the independence of features in determining the class.

4. Literature Review

The authors of [36] grouped the various approaches for sentiment analysis into the following four categories, each with their corresponding challenges:

1) Building Resources (BR) aims at creating lexica, dictionaries, and corpora in which opinion expressions are annotated according to their polarity. The main challenges that confronted the work in this category are the ambiguity of words, multilingually, granularity (the fact that opinions can be expressed in words, sentences, or entire phrases), and the differences in opinion expression among textual genres (blog, newspaper article, microblog). Building resources is not a SA task, but it could help to improve SA.

2) Sentiment Classification (positive and negative and neutral). The main challenges in this category are:

a) the need for contextual analysis of opinion expressions,
b) the importance of contemplating negation and its scope;

c) The detection of irony;

d) The detection of implicit expressions of effect, in objective sentences.

3) Opinion extraction is concerned with finding parts of the text with opinion. The main challenges are detecting the opinion holder and the opinion target.

4) Applications of sentiment analysis to and jointly with other NLP tasks (opinion question answering, opinion summarization, opinion retrieval) or in final integrated systems (e.g., ranking of best products according to the opinions expressed for recommender systems). The main challenges in this area are related to the need to employ high-performance sentiment analysis methods to avoid error propagation and the need to combine different techniques from sentiment analysis with methods from the research area in which it is applied.

However, it is also possible to group the sentiment analysis approaches by linguistic, digital, and possibly hybrid. Some of the major approaches used based on linguistic ones include the below.

In [37], the authors created a code-switching corpus with sentiment annotations. Each word in the corpus is labeled with its language, serving as the starting point to obtain a collection of multilingual tweets. The results show that multilingual models can even outperform the monolingual models on some monolingual sets.

The approach of [38] aims to better predict the sentiment of tweets by using contextual information like (location, timestamp, and author information). They calculated the probability of positive or negative sentiment for each item in the contextual categories. The results are very promising and show that using emoticon-based distant supervision to label the sentiment of tweets is an acceptable method.

Sarcasm is defined as the use of remarks that mean the opposite of what they say, made to hurt someone's feelings or to criticize something humorously. In [39], their approach combines the predictions from two predictors, a contrast-based predictor (that identifies if there is a sentiment contrast within a targeted tweet) and a historical tweet-based predictor (that determines if the sentiment expressed towards an entity in the target tweet agrees with the sentiment expressed by the author towards that entity in the past). They implement four kinds of integrators to combine the predictions and show promising results when applying two predictors.

Their goal in [40] is to create a ranked list of products based on sentiment information. They use different opinion mining methods (dictionary-based, machine learning, comparison-based). The results are promising and encouraging for further research.

Opinion spam can be defined as a fictitious opinion written to sound authentic to mislead the reader. Opinion spam usually is a short text written by an unknown author using a not very well-defined style. These characteristics make the automatic detection of opinion spam a very challenging problem.

The authors [41] introduced a hybrid approach that used three different features: Character n-grams in tokens, Emotion-based feature, LIWC-based feature (Linguistic Inquiry and Word Count). They compared the results with those published previously, and the results were similar with low dimensionality representation for their approach. Then they used Naïve Bayes and SVM algorithms, but they showed only the results of SVM because its performance was the best. They have concluded that character 4-grams in tokens with LIWC variables perform the best using an SVM classifier. That emotions-based feature does not provide much helpful information for detecting deception in reviews.

[42] used a lexicon-based approach to identify sentiment and non-sentiment bearing hashtags by combining existing lexical resources. They developed two classification models, the first using a binary search algorithm to compare each hashtag with each subjective word, and the second using bootstrapping technique. The results show that we can identify non-sentiment hashtags more accurately and precisely than sentiment hashtags.

The authors of [43] presented a preliminary study on different generative models (latent Dirichlet allocation (LDA), Markov chains (MC), and hidden Markov model (HMM)) to generate text with a specific sentiment. Their experiments analyze the scalability, cardinality, and ability to generate text with a sentiment. Sentiment Analysis used (SentiWordNet, Support vector machines (SVMs), Stanford Sentiment Treebank). The results showed that the hidden Markov model achieves a lower F-measure than the Markov chain but can generate a higher number of different texts than Markov chains and achieves a shorter execution time than LDA but greater than MC.

The paper [44] detects the stance categories speculation, contrast, and conditional in English consumer reviews. They presented a comparison between A support vector classifier and a lexicon-based approach. Many researchers evaluated different training data sizes, and the results were comparable to those of previous studies and can be achieved with a smaller amount of training data.

Besides the linguistic aspects, the technical part is also used in some important exercises. The author has applied different algorithms using Python in their studies. (45) (The earlier reviews of the sentiment analysis are carried out by Naaima Boudad et al. in 2018 (46), by Ghallab in 2020(47), and by Oueslati in 2020 (48). Recently the deep learning approach is deployed by Ayah Soufan (49).

The Arabic Language:
Arabic is considered one of the top 10 languages on the Internet-based on the ranking carried out in 2016, and hundreds of millions speak it. The Arabic language is a Semitic language with rich morphology. Classical, Modern Standard Arabic (MSA), and colloquial are the three main variants. The Arabic language is written from right to left and consists of 28 letters.

Therefore, there are many challenges facing sentiment analysis for Arabic text, such as:

• Morphological complexities and dialectal varieties of the Arabic language require advanced pre-processing and lexicon-building steps beyond what is applicable for the English language domain. [14] [15]

• Different words with different meanings can be produced using the same three-letter root.

• The absence of rigid and strict rules in adding punctuation in MSA text makes it very hard to identify the sentence boundaries. [46]

• There is also no capitalization in Arabic, which determines sentence boundaries a crucial and challenging task for NLP in the Arabic language, especially for sentiment analysis.

4.1. Arabic Sentiment Corpora
There exist limited annotated resources for sentiment analysis, such as labeled corpora and polarity lexica in Arabic. Most of these resources are manually annotated at either the document or sentence level.

In [47], the authors proposed (*ArSenL*) the first publicly available lexicon as a large-scale Arabic Sentiment Lexi con. *ArSenL* is constructed using a combination of English *SentiWordne*t (ESWN), Arabic WordNet, and the Arabic Morphological Analyzer (*AraMorph*).

Another publicly available large-scale Standard Arabic Sentiment Lexicon (SLSA) was proposed by [48]. The construction of SLSA is based on a combination of the lexicon of *AraMorph* with *SentiWordNet*.

In [49], the authors have generated the Opinion Corpus for Arabic (OCA), the first corpus created for the Arabic language. Then in [50], they have used an automatic machine translation tool to generate the English Version of OCA (EVOCA). Both corpora are freely available for the research community and include 500 reviews obtained from specialized Arabic web pages related to movies and films, 250 positives and 250 negatives.

AWATIF is a multi-genre and manually built corpus for Arabic Subjectivity, and SA was proposed in [51]. AWATIF is extracted from three different resources:

- The Penn Arabic Treebank (PATB)
- Wikipedia user talk pages
- Conversation threads from web forums of seven other sites

The authors of [52] presented LABR, the Large-scale Arabic Book Review dataset, consisting of over 63,000 book reviews, each with a rating of 1 to 5 stars. The classifiers (MNB, BNB, SVM, Passive, Aggressive, SGD, Logistic Regression, Linear Perceptron, KNN) used in this experiment are widely used in sentiment analysis. They can be considered a baseline benchmark for any further investigations on the dataset.

The authors of [53] manually created a flexible and relatively extensive standard Arabic corpus for SA and opinion mining for Arabic reviews and comments. This corpus is freely available and consists of 1,442 reviews covering 250 topics distributed equally among five domains: Economy, Food-Life style, Religion, Sport, and Technology. In addition to MSA, this corpus contains reviews written in the five main Arabic dialects (Egyptian, Levantine, Arabian Peninsula, Mesopotamian, and Maghrebi group). It is flexible, where users can add, delete or revise it.

4.2. Arabic Sentiment Analysis

Here, we introduce an overview of recent important papers which study sentiment analysis for the Arabic language.

In [54], they investigated the effectiveness of the Multi-Way Sentiment Analysis (MWSA) approach for Arabic reviews. MWSA focuses on sentiments conveyed through a rating or scoring system. They proposed to use hierarchical classification. For example, one classifier can be trained trained to distinguish between positive and negative reviews while another classifier is trained independently to distinguish between strong and weak positive reviews. Similarly, a third classifier is trained separately to differentiate between strong and weak negative reviews. They presented two different hierarchical structures and compared their accuracy with the flat structure. The results showed that, in general, hierarchical classifiers gave significant improvements over flat classifiers.

A comparison between two classification techniques, SVM and NB, uses Arabic political tweets were presented in [55]. The results showed that the Naïve Bayesian method is the highest accuracy and the lowest error rate.

[56] built a sentiment lexicon with about 120,000 Arabic terms. The process is divided into three steps: collect Arabic stems, translate them into English, and use online English sentiment lexicons to determine the sentiment value of each word. They compared their approach with the keyword-based approach, and the proposed approach showed excellent results.

The authors of [57] defined rules to capture the morphology of negations in Arabic. [58] proposed a SA and Negation Resolving system for Arabic Text Entailment. The results showed that resolving the negation and classifying text polarity increases the performance of detecting the entailment relation and non-entailment relation.

In [59], the authors introduce a semi-supervised approach for sentiment analysis using a high coverage Arabic sentiment words lexicon (ArSeLEX) which is manually created and automatically expandable, and sentiment Arabic idioms/saying phrase lexicon to improve the classification process. They built a corpus containing 2000 Arabic sentiment statements, including 1000 MSA tweets, Arabic dialect tweets, and 1000 microblogs. Then, they used a support vector machine (SVM) classifier with linguistically, syntactically, novels and rich feature sets to handle the valence shifters (negation, intensifiers), question, and supplication terms. The results obtained show that their approach is very promising.

In [60], the authors have presented a light lexicon-based mobile application for sentiment mining of Arabic Tweets. The mobile application was designed to minimize the energy consumption of the mobile by having an algorithm with minimal computational needs and no remote communication for computation. The authors used decision trees as a classification model to classify tweets into positive, negative, and objective.

The first study to measure the impact of automatically translated data on the accuracy of sentiment analysis of Arabic tweets was proposed by [61]. They investigate using Machine Translation tools to perform SA for under-resourced languages, such as Arabic, an effective and efficient alternative to building SA classifiers from scratch. At first, they retrieve Arabic tweets from the Twitter public stream. They translated the dataset by using Google Translate and Microsoft Translator Service. Then, they used the Stanford Sentiment Classifier (SSC) to automatically assign sentiment labels (positive, negative) to translated tweets. After that, they compared the performance of their approach to standard SA approaches (lexicon-based, supervised, and distant supervision approaches), and the results showed comparable performance.

This paper performs SA on Arabic reviews and comments using four machine learning algorithms: Support Vector Machine (SVM), Back-Propagation Neural Networks (BPNN), Naïve Bayes, and Decision Tree. The authors of [62] collected 2000 Arabic reviews from social media to evaluate the different machine learning algorithms based on sentiment analysis. Results showed that the SVM classifier achieves the highest accuracy rate compared with other classifiers.

5. Related Topics:Contact

5.1. Deep learning:

In recent years, deep learning models have achieved remarkable results in computer vision [74] and speech recognition [75].

According to [76], understanding the compositionality in sentiment detection requires more prosperous supervised training and evaluation resources and more powerful composition models. To handle this, they introduce a Stanford Sentiment Treebank, the first corpus with fully labeled parse trees that allows for a complete analysis of the compositional effects of sentiment in language. It includes fine-grained sentiment labels for 215,154 phrases in the parse trees of 11,855 sentences. Then, they proposed a new model called the Recursive Neural Tensor Network (RNTN), which takes as input phrases of any length and represents a phrase through word vectors and a parse tree and then computes vectors for higher nodes in the tree using the same tensor-based composition function. They compared their model to several supervised, compositional models such as standard recursive neural networks (RNN) [77], matrix-vector RNNs [78], Naive Bayes (NB), bi-gram NB, and SVM. The results showed that this model significantly outperforms previous models and captures negation of different sentiments and scope more accurately than previous models. Also, the authors observe that the recursive models work well on the shorter text while BOW features with NB and SVM perform well only on longer sentences.

This classifier was used in sentiment analysis for Arabic tweets [61], which we mentioned above.

Convolutional neural networks (CNN) utilize layers with convolving filters applied to local features [79].

A series of experiments with convolutional neural networks (CNN) were introduced in [80]. They prepared a simple CNN with one layer of convolution on top of word vectors obtained from an unsupervised neural language model. They used the publicly available word2vec vectors trained on 100 billion words from Google News. They showed that a simple CNN with little hyper-parameter tuning and static vectors achieves excellent results on multiple benchmarks.

A deep learning-based approach to aspect-based sentiment analysis, which employs a convolutional neural network for aspect extraction and sentiment analysis, was presented in [82]. They reported convincing results for multilingual aspect-based sentiment analysis.

In [83], they introduced a generalization of the standard LSTM architecture to tree-structured network topologies and showed its superiority for representing sentence meaning over a sequential LSTM. They evaluated the Tree-LSTM architecture on two tasks: semantic relatedness prediction on sentence pairs and sentiment classification of sentences. Their experiments showed that Tree-LSTMs outperform both tasks' existing systems and sequential LSTM baselines.

The first work investigates the merit of using deep models for sentiment analysis in Arabic proposed in [84]. They explore four deep learning models:

- Deep neural networks (DNN)
- Deep Belief Networks (DBN)
- Deep Auto Encoder (DAE)
- Combined DAE with DBN

They used the Linguistic Data Consortium Arabic Tree Bank (LDC ATB) dataset to evaluate the proposed models. The input data to the first three models depends on the BoW model, utilizing *ArSenL* lexicon scores [47]. For the fourth model (RAE), the input is the raw word indices that constitute each sentence. The RAE model outperforms all the other models and over the best-reported results in the literature on the same LDC ATB dataset in the sentiment classification task for Arabic. Semantic context and the parsing order of words are considered in this model. At the same time, no lexicon is used, and no special features are used, but only raw words as input.

5.2. Multimodal Sentiment Analysis

New sources of opinion mining and sentiment analysis abound. Webcams, smartphones, touchpads, or other devices let users post opinions in an audio or audiovisual format rather than in text. Multimedia content has become more prevalent in social networks, especially on Tumblr, Instagram, and Flickr platforms.

In [85], built large-scale Visual Sentiment Ontology based on psychological theories and web mining and trained detectors of selected visual concepts for sentiment analysis.

In [86] and [87], the concept of transfer learning Deep CNN from large-scale image classification applied to the problem of sentiment prediction.

There is almost no research that focuses on multimodal sentiment and opinion analysis. The authors [88] combine acoustic, textual, and video features to assess opinion polarity in a new dataset consisting of 47 YouTube videos. They demonstrate significant improvement in left-one-video-out evaluation using Hidden Markov Models for classification.

Also, the authors of [89] have presented a multimodal sentiment analysis framework, which includes sets of relevant features for text and visual data. Several supervised classifiers, namely Naïve Bayes, SVM, ELM, and Neural Networks, were employed on the fused feature vector to obtain the sentiment of each video segment.

All research relied on transcripts to analyze the text and not the actual spoken word.

6. Discussion and Analysis

Table 1 summarizes the most critical and recent papers that deal with Arabic sentiment analysis. We aggregated this data from three surveys presented in [90] [91] [13], in addition to some recent papers. Table 1 contains the article reference in the first column and the year of publishing paper in the next one. The objectives of the articles are illustrated in the third column. They are divided into Sentiment Analysis in general (SA), Building Resource (BR), and Sentiment Classification (SC). In the fourth column, Domain Oriented means that domain-specific data are used in the SA process. The fifth column shows the used algorithms. The sixth column specifies whether the article uses SA techniques for general Analysis of Text (G) or solves the problem of binary classification (Positive/Negative). The seventh and eighth columns illustrate the scope (reviews, news articles, web forums, tweets) and the type (Modern Standard Arabic, Colloquial Arabic) of the data used for evaluating the article's algorithms. The last column specifies the level of sentiment analysis (Document, Sentence, and Aspect).

We can observe that the work on Arabic SA is very little, but the number of papers is growing every year. In figure 1, we can note how big the gap is between research that has been conducted in Arabic sentiment analysis and English sentiment analysis. Data in figure 1 is collected by using relevant keywords in the sentiment analysis field from the Google Scholar website for a specific period.

The Machine learning-based approach was used in most articles because of its simplicity and ability to learn from features for classification automatically. Most features are challenging to use by a lexicon-based method. It's more frequent because of its domain independence. Hybrid methods are limited because of their computational complexity.

As discussed above, the datasets to evaluate the algorithm can be reviews, web pages, social network comments, etc. Most of the works were on reviews, as we see in figure 3, and recently the trend is to evaluate tweets and social networks comments. The tweet is a message composed of 140 characters long and has hashtags as a unique feature.

From figure 4, we can see that works on aspect level for Arabic sentiment analysis are rare even though it is essential to discover sentiments on entities and their aspects.

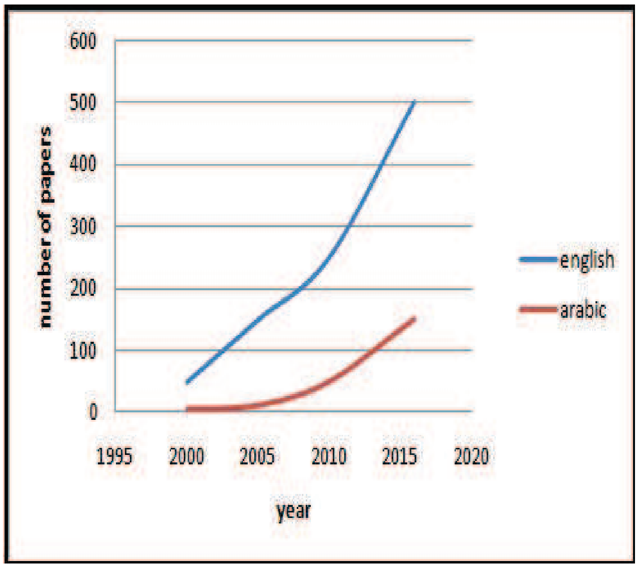

Figure 1. Number of papers for Arabic and English SA

7. Conclusions and Future Work

In this paper, we introduce a general overview of sentiment analysis and present a survey of Arabic SA. We categorized and summarized these articles that give contributions to many sentiment analysis-related fields.

After analyzing these articles, we can conclude that sentiment analysis for English language research has advanced. There is a big gap between it and sentiment study for the Arabic language. Also, the improvement of Sentiment Classification and Feature Selection algorithms is still an open field for research. We observe that Naïve Bayes and Support Vector Machines are the most frequently used ML algorithms for solving Sentiment Classification problems. Many enhancements could be done in Arabic Sentiment Analysis, such as research-based on hybrid approaches and deals with negations, spam, sarcasm, and aspect level. It is important to consider the media information and the context of the text.

Therefore, we believe that combining semantic web technology and sentiment analysis could be valuable. Also, we are looking to take advantage of deep learning in this field.

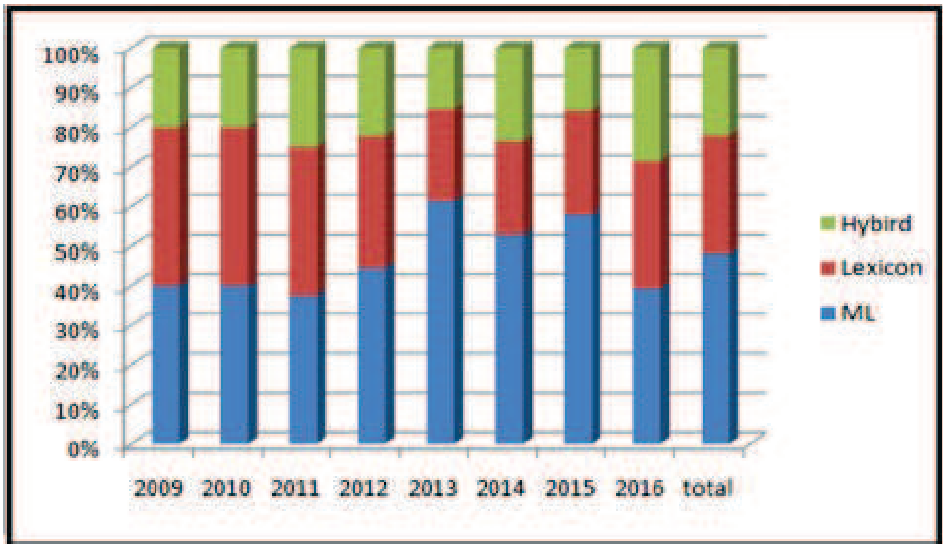

Figure 2. Percentage of articles according to the algorithmic approach over years

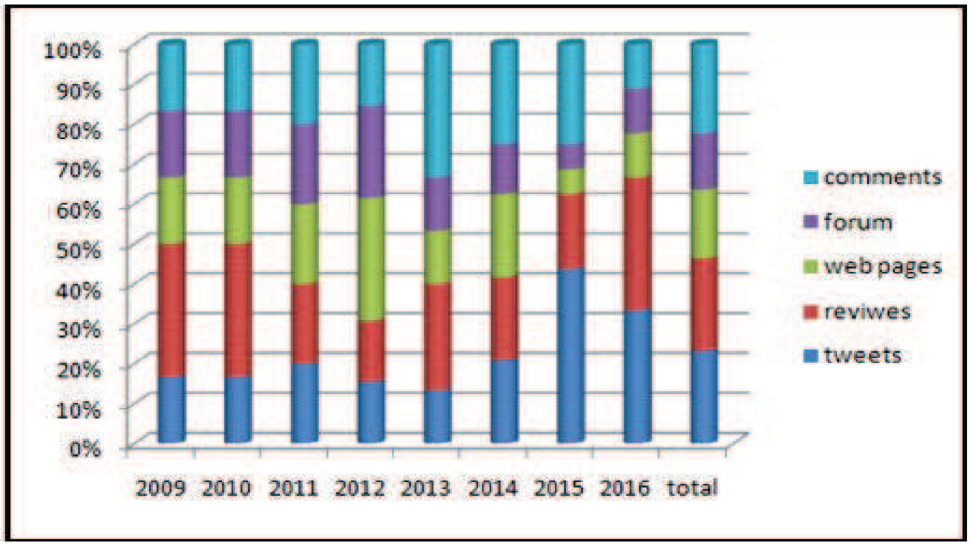

Figure 3. Percentage of articles according to the dataset scope over years

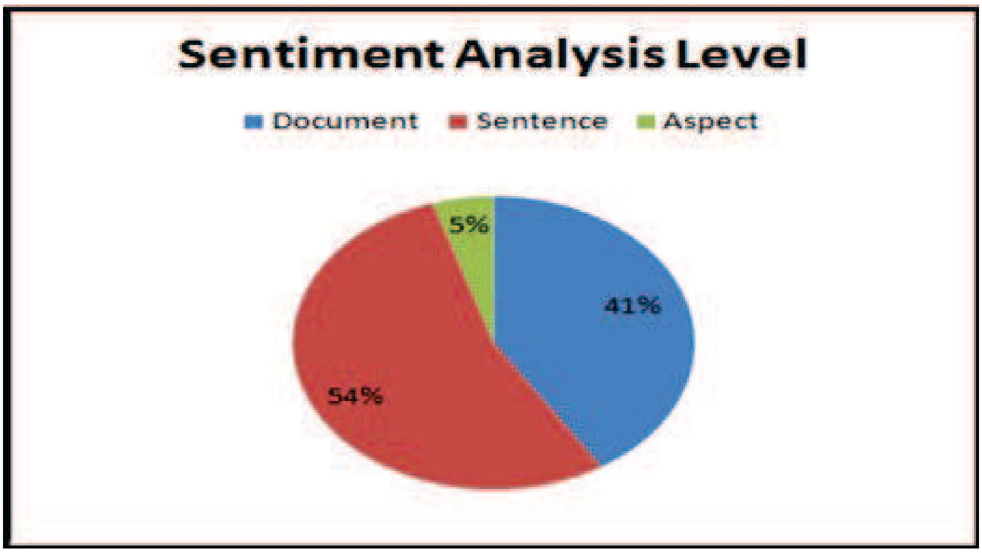

Figure 4. Percentage of articles according to the SA level

Ref.	Year	Task	Domain Oriented	Algorithm Used	Polarity	Data Scope	Type	SA level
[7]	2016	SA	YES	SVM/NB	Pos/Neg/ Neutral	Arabic Tweets	MSA	Document
[9]	2010	SC	NO	Lexicon-Based	Pos/ Neg	Arabic Movies	MSA	Document/Sentence
[10]	2012	SC	YES	SVM/NB	Pos/Neg	Arabic Tweets	MSA/Dialects	Sentence
[11]	2016	SA	YES	Ontology-based	Pos/ Neg	Hotel Reviews	MSA	Aspect
[14]	2014	SA	NO	SVM/NB/KNN/ DT	Pos/Neg	Reviews	MSA/Dialects	Sentence
[15]	2013	SA	YES	SVM/NB	Pos/Neg	Social comments	MSA/ Egyptian Dialects	Sentence
[16]	2013	SA	YES	SVM	G/Pos/Neg/Netural	Reviews/ Newswire	MSA/ Dialectats	Sentence
[17]	2012	SA	NO	SVM	Pos/Neg	Twitter/ Social Comments	MSA/Egyptian Dialect	Sentence
[24]	2014	SA	YES	SVM/NB	Pos/Neg	Reviews	MSA	Sentence
[25]	2014	SA	NO	SVM/KNN/NB	Pos/Neg	Arabic Tweets/ Comments	MSA/Dialects	Sentence
[26]	2014	SC	NO	ML	Pos/Neg/ Neutral	Arabic Tweets	MSA/Jordanian Dialect/ Arabizi	Document
[27]	2012	SA	NO	Hybird	Pos/ Neg/ Neutral	Twitter/Wikipedia pages/Maktoob chat/ Web forum	MSA/Arabic	Sentence
[28]	2013	SC	YES	SVM/NB/DT	Supportive/Attacking/ Neutral	Facebook Comments	MSA/ Egyptian Dialect	Sentence
[29]	2013	SC	YES	ML	Pos/Neg	Arabic news comments	Arabic Dialects	Document
[30]	2013	SA	NO	ML/Lexicon-based	Pos/Neg	Arabic Tweets	MSA/Dialects	Sentence
[31]	2014	SC	YES	Lexicon-Based	G / Pos/Neg	Social and News Sites	MSA/Dialects	Document
[32]	2014	SC	NO	Hybrid	G / Pos/Neg	Twitter	MSA/Egyptian Dialect	Sentence
[33]	2015	SA	YES	ML	Pos/Neg	Arabic Tweets	MSA/Dialects	Document
[34]	2014	SA	YES	Hybrid	Pos/Neg	Comments	MSA/ Dialects	Sentence
[35]	2012	SC	NO	ML	Pos/Neg	Arabic Facebook posts	Syrian, Egyptian, Iraqi and Lebanese Dialects	Document
[46]	2013	SA	NO	Rule-based classifiers	Pos/Neg	Arabic movie reviews	MSA	Sentence
[47]	2014	BR	NO	SVM/NB	G/Pos/Neg	Reviews	MSA	Document
[50]	2011	BR	NO	SVM/NB	Pos/ Neg	Movies Reviews	Arabic/English	Sentence
[51]	2012	BR	NO	Lexicon-based	Pos/ Neg/ Neutral	Wikipedia/Web forums/	MSA	Sentence
[52]	2014	BR	YES	SVM/NB/KNN/ Logistic Regression	Pos/Neg/ Neutral	Book Reviews	MSA/Dialects	Sentence
[53]	2016	BR	YES	SVM/NB	Pos/neg/ Neutral/ Spam	Maktoob Yahoo! Website	MSA/Dialects	N/A
[54]	2016	SC	NO	SVM/NB/KNN/ DT	(1 TO 5) Rating Classes	Book Reviews	MSA/Dialects	Sentence
[55]	2016	SC	YES	SVM/NB	Pos/Neg	Arabic Tweets	MSA	Sentence
[56]	2015	SA	NO	Lexicon-Based	Pos/Neg	Arabic Tweets	MSA	Sentence
[57]	2015	SA	NO	Lexicon-Based	Pos/Neg	Arabic Reviews	MSA	Sentence
[58]	2015	SC	Yes	Lexicon-Based	Pos/Neg	Arabic Text	MSA/Dialects	Document
[59]	2015	SA	NO	Hybird	Pos/Neg	Tweets/Reviews/ Comments	MSA/ Egyptian Dialects	Aspect
[60]	2015	SA	NO	Lexicon-Based	Pos/Neg	Arabic Tweets	MSA	Document
[61]	2015	SC	NO	MT-based	Pos/Neg	Arabic Tweets	MSA/ Dialectal Arabic	Sentence
[63]	2013	SA	YES	SVM/NB/KNN	Spam/Not Spam	online Arabic economic websites	MSA	Document
[68]	2016	SA	YES	Ontology-based	Pos/Neg	Arabic Tweets	MSA/ Jordanian Dialect	Sentence
[90]	2014		NO	Hybird	Pos/Neg/ Neutral	Arabic Tweets	MSA	Sentence

Table 1. A summary for most important and recent papers which deal Arabic sentiment analysis

References

[1] Financial Times. (2016) Chief's 'hubris' steered Cambridge Analytica to data scandal. [Online]. https://www.ft.com/content/e4e95b6c-2dac-11e8-9b4b-bc4b9f08f381

[2] Bo., Lee., Lillian and others Pang, *Opinion mining and sentiment analysis*.: Foundations and Trends

[3] Binali, Haji ., Potdar, Vidyasagar., Wu, Chen. (2009). A state of the art opinion mining and its application domains, *In*: *I*ndustrial Technology, 2009. ICIT 2009. *IEEE International Conference on*, 1-6.

[4] Liu, Bing. (2012). Sentiment analysis and opinion mining, *Synthesis lectures on human language technologies*, volume 5, 1-167.

[5] Liu, Bing. (2015). *Sentiment analysis: Mining opinions, sentiments, and emotions*. Cambridge University Press, 2015.

[6] Indurkhya, eds., Damerau, Fred J. (2010). Handbook of natural language processing, *In*: Chapman and Hall/ CRC.

[7] Al-Rubaiee, Hamed., Qiu, Renxi., Li, Dayou., Identifying Mubasher software products through sentiment

analysis of Arabic tweets, *In*: *I*ndustrial Informatics and Computer Systems (CIICS), 2016 International Conference on, 2016, 1-6.

[8] Bruce, Rebecca F., Wiebe, Janyce M. (1999). Recognizing subjectivity: a case study in manual tagging, *Natural Language Engineering*, 187-205.

[9] Farra, Noura., Challita, Elie., Assi, Rawad Abou., Hajj, Hazem. (2010). Sentence-level and document-level sentiment mining for arabic texts, *In*: *Data Mining Workshops (ICDMW), 2010 IEEE International Conference on*, 2010, 1114-1119.

[10] Shoukry, Amira., Rafea, Ahmed. (2012). Sentence-level Arabic sentiment analysis, *In*: Collaboration Technologies and Systems (CTS), 2012 International Conference on, 2012, 546-550.

[11] Alkadri, Abdullah M., ElKorany, Abeer M. (2016). Semantic Feature Based Arabic Opinion Mining Using Ontology, *INTERNATIONAL JOURNAL OF ADVANCED COMPUTER SCIENCE AND APPLICATIONS*, Volume 7, p. 577-583, 2016.

[12] Kontopoulos, Efstratios., Berberidis, Christos., Dergiades, Theologos., Bassiliades, Nick. (2013). Ontology-based sentiment analysis of twitter posts, *Expert systems with applications*, volume 40, p. 4065-4074.

[13] Assiri, Ahmed Emam., Aldossari, Hmood. (2015). Arabic Sentiment Analysis: A Survey, in *International Journal of Advanced Computer Science & Applications, 1.6*, 2015, 75-85.

[14] Bao, Yanwei., Quan, Changqin., Wang, Lijuan., Ren, Fuji. (2014). The role of pre-processing in twitter sentiment analysis, *In*: International Conference on Intelligent Computing, 615-624.

[15] Haddi, Emma., Liu, Xiaohui., Shi, Yong. (2016). The role of text pre-processing in sentiment analysis, *Procedia Computer Science*, volume 17, 26-32.

[16] Oraby, Shereen., El-Sonbaty, Yasser., El-Nasr, Mohamad Abou. (2013). Exploring the Effects of Word Roots for Arabic Sentiment Analysis, in *IJCNLP*, 471-479.

[17] Amira., Rafea., Shoukry, Ahmed. (2012). Preprocessing Egyptian dialect tweets for sentiment mining, *In*: The Fourth Workshop on Computational Approaches to Arabic Script-based Languages, 2012, p. 47.

[18] Kummer, Olena., Savoy, Jacques., Argand, Rue Emile. (2012). Feature selection in sentiment analysis.

[19] Pang, Bo., Lee, Lillian., Vaithyanathan, Shivakumar. (2002). Thumbs up?: sentiment classification using machine learning techniques, *In*: Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10, 79–86.

[20] Mejova, Yelena., Srinivasan, Padmini. (2011). Exploring Feature Definition and Selection for Sentiment Classifiers., *In*: *ICWSM*.

[21] Wilson, Derek., et al. (2009). SUPERFAMILY—sophisticated comparative genomics, data mining, visualization and phylogeny, *Nucleic Acids Research, Volume 37, Issue suppl_1*, D380–D386, 1 January 2009.

[22] Aggarwal, Charu C., Zhai, ChengXiang. (2012). A survey of text classification algorithms, *In*: *Mining text data*.: Springer, p 163-222.

[23] Agarwal, Basant., Mittal, Namita. (2012).Categorical probability proportion difference (CPPD): a feature selection method for sentiment classification, *In*: Proceedings of the 2nd workshop on sentiment analysis where AI meets psychology, COLING, p 17-26.

[24] Elawady, Rasheed M., Barakat, Sherif., Elrashidy, Nora M. (2014). Different feature selection for sentiment classification, *International Journal of Information Science and Intelligent System*, Voulme 3, 137-150.

[25] Rehab, M., Qarqaz, Islam Duwairi. (2014). Arabic sentiment analysis using supervised classification, in *Future Internet of Things and Cloud (FiCloud), 2014 International Conference on*, 2014, 579—583.

[26] RM., Marji., Raed., Sha'ban., Narmeen., Rushaidat., Sally Duwairi. (2014). Sentiment analysis in arabic tweets, in *Information and communication systems (icics), 2014 5th international conference on*, 2014, p1—6.

[27] Muhammad., Mona, Diab., SAMAR, K. (2012).: Subjectivity and sentiment analysis for Arabic social media.

[28] Ali, Alaa El-Dine., El-taher., Hamouda, Fatma El-zahraa. (2013). Sentiment analyzer for arabic comments system, *Int. J. Adv. Comput. Sci. Appl*.

[29] Rahman, Abdel., Hedar, Doss, M. (2013). Mining social networks arabic slang comments, *In*:: IEEE Symposium on Computational Intelligence and Data Mining (CIDM).

[30] Nawaf, A., Ahmed., Nizar, A., Shehab, Mohammed A ., Al-Ayyoub., Mahmoud Abdulla. (2013). Arabic sentiment analysis: Lexicon-based and corpus-based, *In*: Applied Electrical Engineering and Computing Technologies (AEECT), 2013 IEEE Jordan Conference on, 1-6.

[31] Mohammed, N., Gigieh, Amal H., Alsmadi, Izzat M., and Wahsheh, Heider A., Haidar, Mohamad M Al-Kabi., Opinion mining and analysis for arabic language, *International Journal of Advanced Computer Science and Applications (IJACSA), SAI Publisher*, 5 (5) 181—195.

[32] Nagwa., Nagi., Khaled., El-Ebshihy., Alaa and Apady., Esraa and Hafez., Omneya and Mostafa.,

[33] Nagwa and Nagi., Khaled., El-Ebshihy., Alaa., Apady., Esraa., Hafez., Omneya., Mostafa., Samar., Ibrahim, Shimaa El-Makky. (2014) Sentiment analysis of colloquial Arabic tweets, *In*: ASE BigData/SocialInformatics/PASSAT/BioMedCom 2014 Conference, Harvard University.

[34] Rawan, T., Wahsheh, Heider A., Alsmadi, Izzat M., Al-Kabi, Khasawneh, Mohammed N. (2015). Arabic sentiment polarity identification using a hybrid approach, in

Information and Communication Systems (ICICS), 2015 6th International Conference, p 148-153.

[35] Khalid and Omar., Nazlia Khalifa. (2014). A hybrid method using lexicon-based approach and naive Bayes classifier for Arabic opinion question answering, *Journal of Computer Science*, 10 (10) 1961-1968.

[36] Maher, M., Zantout, Rached N., Hamandi, Lama., Elkabani, Islam Itani. (2012). Classifying sentiment in arabic social networks: Naive search versus naive bayes, *In*: Advances in Computational Tools for Engineering Applications (ACTEA), 2012 2nd International Conference on, 2012, 92-197.

[37] Andr, Patricio Mart., Alexandra Balahur. (2012). Subjectivity and sentiment analysis: An overview of the current state of the area and envisaged developments.

[38] David Vilares., Alonso, Miguel A., Carlos, G. (2015). Sentiment analysis on monolingual, multilingual and code-switching twitter corpora, *In*: 6TH WORKSHOP ON COMPUTATIONAL APPROACHES TO SUBJECTIVITY, SENTIMENT AND SOCIAL MEDIA ANALYSIS WASSA 2015, 2015, p. 2.

[39] Soroush, Vosoughi., Zhou, Helen., Roy, Deb. (2016). Enhanced twitter sentiment classification using contextual information, *arXiv preprint arXiv:1605.05195*, 2016.

[40] Khattri, Anupam., Joshi, Aditya., Bhattacharyya, Pushpak., Carman, Mark James. (2015). Your Sentiment Precedes You: Using an author_s historical tweets to predict sarcasm, *In: 6th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis (WASSA)*, 2015, p. 25.

[41] Kessler, Wiltrud., Klinger, Roman., Kuhn, Jonas. (2015). Towards Opinion Mining from Reviews for the Prediction of Product Rankings, in *6TH WORKSHOP ON COMPUTATIONAL APPROACHES TO SUBJECTIVITY, SENTIMENT AND SOCIAL MEDIA ANALYSIS WASSA 2015*, p. 51.

[42] Cagnina, Leticia C., Rosso, Paolo. (2015). Classification of deceptive opinions using a low dimensionality representation, in *6TH WORKSHOP ON COMPUTATIONAL APPROACHES TO SUBJECTIVITY, SENTIMENT AND SOCIAL MEDIA ANALYSIS WASSA 2015*, 2015, p. 58.

[43] Simeon, Credell., Hilderman, Robert. (2015). Using combined lexical resources to identify hashtag types, in *6TH WORKSHOP ON COMPUTATIONAL APPROACHES TO SUBJECTIVITY, SENTIMENT AND SOCIAL MEDIA ANALYSIS WASSA 2015*, p. 169.

[43] Maqsud, Umar. (2015). Synthetic Text Generation for Sentiment Analysis, *In*: 6TH WORKSHOP ON COMPUTATIONAL APPROACHES TO SUBJECTIVITY, SENTIMENT AND SOCIAL MEDIA ANALYSIS WASSA 2015, 2015, p. 156.

[44] Skeppstedt, Maria., Schamp-Bjerede, Teri., Carita, Magnus Sahlgren. (2015). Paradise Andreas Kerren. (2015). Detecting speculations, contrasts and conditionals in consumer reviews, *In:* Proceedings of the 6th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, p. 162-168.

[45] Folorunsho, Dhikrullah. (2020). Arabic Sentiment Analysis. Data Science-Oct 13.

[46] Rdouan, Naaima Boudad., Faizi Rachid Oulad Haj Thami Raddouane Chiheb. (2018). Sentiment analysis in Arabic: A review of the literature, *Ain Shams Engineering Journal*. 9 (4) 2479-2490)

[47] Ghallab, Abdullatif., Mohsen, Abdulqader., Ali, Yousef. (2020). Arabic Sentiment Analysis: *A Systematic Literature Review. Applied Computational Intelligence and So Computing*. Volume 2020, Article ID 7403128, p. 21.

[48] Oueslati, Oumaima., Cambria, Erik., HajHmida, Ben., Ounelli, Habib. (2020). A Review of Sentiment Analysis, Research in Arabic Language. p. 1-46. www.internetworldstats.com/stats7.htm)

[49] Soufan, Ayah. (2019). Deep Learning for Sentiment Analysis of Arabic Text, *In:* ArabWIC 2019: Proceedings of the ArabWIC 6th Annual International Conference Research Track. March 2019 Article No. 20. 1–8[50] Harrag, Fouzi. (2014). Estimating the Sentiment of Arabic Social Media Contents: A Survey, *Research Center of College of Computer and Information Sciences, King Saud University under the project code (RC131001)*, 2014.

[51] Oraby, Shereen., El-Sonbaty, Yasser., El-Nasr, Mohamad Abou. (2013). Finding opinion strength using rule-based parsing for arabic sentiment analysis, in *Mexican International Conference on Artificial Intelligence*, 2013, 509-520.

[52] Badaro, Gilbert., Baly, Ramy., Hajj, Hazem., Habash, Nizar., El-Hajj, Wassim. (2014). A large scale Arabic sentiment lexicon for Arabic opinion mining, *ANLP 2014*, volume 165.

[53] Eskander, Ramy., Rambow, Owen. (2015). SLSA: A Sentiment Lexicon for Standard Arabic., *In*: EMNLP, 2015, p. 2545-2550.

[54] Rushdi-Saleh, Mohammed., Mart, Teresa M., Alfonso Ure, L., Jos, (2011). OCA: Opinion corpus for Arabic, *Journal of the American Society for Information Science and Technology*, 62, 2045-2054.

[55] Mohammed, Rushdi-Saleh., Maria, Mart, Teresa., Ure, L. Alfonso., Jos. (2011). Bilingual experiments with an arabic-english corpus for opinion mining.

[56] Abdul-Mageed, Muhammad., Mona T, Diab. (2012). AWATIF: A Multi-Genre Corpus for Modern Standard Arabic Subjectivity and Sentiment Analysis., *In*: LREC, 2012, p. 3907-3914.

[57] Nabil, Mahmoud., Aly, Mohamed., Atiya, Amir. (2014). LABR: A Large Scale Arabic Sentiment Analysis Benchmark, *arXiv preprint arXiv:1411.6718*.

[58] Al-Kabi, Mohammed., Al-Ayyoub, Mahmoud.,

Alsmadi, Izzat., Wahsheh, Heider. (2016). A prototype for a standard arabic sentiment analysis corpus., *Int. Arab J. Inf. Technol.*, 13, 163-170.

[59] Al-Ayyoub, Mahmoud., Nuseir, Aya., Kanaan, Ghassan., Al-Shalabi, Riyad. (2016). Hierarchical classifiers for multi-way sentiment analysis of arabic reviews, *International Journal of Advanced Computer Science and Applications (IJACSA)*, volume 7, 531-539.

[60] Elghazaly, Tarek., Mahmoud, Amal., Hefny, Hesham., A. (2016). Political sentiment analysis using twitter data,*In: Proceedings of the International Conference on Internet of things and Cloud Computing*, 2016, p. 11.

[61] Al-Ayyoub, Mahmoud., Essa, Safa Bani., Alsmadi, Izzat. (2015). Lexicon-based sentiment analysis of arabic tweets, *International Journal of Social Network Mining*, volume 2, 101-114.

[62] Duwairi, Rehab M., Alshboul, Mohammad A. (2015). Negation-aware Framework for Sentiment Analysis in Arabic Reviews, *In*: *F*uture Internet of Things and Cloud (FiCloud), 2015 3rd International Conference on, 2015, p. 731-735.

[63] AL-Khawaldeh, Fatima T. (2015). A Study of the Effect of Resolving Negation and Sentiment Analysis in Recognizing Text Entailment for Arabic, *World of Computer Science and Information Technology Journal* (WCSIT), volume 5, p. 124-128.

[64] Ibrahim, Hossam S., Abdou, Sherif M., Gheith, Mervat. (2015). Sentiment analysis for modern standard arabic and colloquial, *arXiv preprint arXiv:1505.03105*, 2015.

[65] Gilbert Badaro., et al. (2015). A light lexicon-based mobile application for sentiment mining of Arabic tweets, in *ANLP Workshop,* 2015, p. 18.

[66] Refaee, Eshrag., Rieser, Verena. (2015). Benchmarking Machine Translated Sentiment Analysis for Arabic Tweets., *In*: *HLT-NAACL*, 2015, 71-78.

[67] Hammad, Mustafa., Al-awadi, Mouhammd. (2016). Sentiment Analysis for Arabic Reviews in Social Networks Using Machine Learning, in *Information Technology: New Generations*.: Springer, 2016, p. 131-139.

[68] Hammad, Ahmad S. J. Abu. (2013). An approach for detecting spam in arabic opinion reviews, Islamic University of Gaza, Ph.D. dissertation 2013.

[69] Ruidi., Benjamins., Richard, V., Fensel., Dieter Studer. (1998). Knowledge engineering: principles and methods, *Elsevier*, 161-197.

[70] Li, Sheng-Tun., Tsai, Fu-Ching. (2013). A fuzzy conceptualization model for text mining with application - in opinion polarity classification, *Knowledge-Based Systems*, 39, 23-33, 2013.

[71] Rajagopal, Hari. (2005). JENA: A Java API for ontology management, *IBM Corporation, Colorado Software Summit*, p. 23-28, 2005.

[72] Alfonso, Ms Meleesa., Sardinha, Ms Razia. INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY ONTOLOGY BASED ASPECT LEVEL OPINION MINING".

[73] Tartir, Samir., Abdul-Nabi, Ibrahim. (2016). Semantic Sentiment Analysis in Arabic Social Media, *Journal of King Saud University-Computer and Information Sciences*.

[74] AlAsmar, Ahmed M. I. (2016). Feature Based Approach in Arabic Opinion Mining Using Ontology, The Islamic University-Gaza, Ph.D. dissertation.

[75] Robert., Chin., Joshua., Havasi., Catherine Speer. (2017). ConceptNet 5.5: *An Open Multilingual Graph of General Knowledge*. *AAAI*, p. 4444-4451.

[76] Christiane, Fellbaum. (2010). Theory and applications of ontology: computer applications, *Springer*, 231-243.

[77] Marie-Catherine., Manning., De Marneffe, Christopher D. (2008). Stanford typed dependencies manual, Stanford University, Technical report.

[78] Gilbert., Baly., Ramy., Hajj., Hazem., Habash., Nizar., El-Hajj., Wassim Badaro. (2014). A large scale Arabic sentiment lexicon for Arabic opinion mining, *In*: *Proceedings of the EMNLP 2014 Workshop on Arabic Natural Lan guage Processing (ANLP)*, 165—173.

[79] Alex., Sutskever., Ilya., Hinton., Krizhevsky, Geoffrey E. (2012). Imagenet classification with deep convolutional neural networks, *Advances in neural information processing systems*, 1097-1105, 2012.

[80] Alex., Mohamed., Abdel-rahman., Hinton., Geoffrey Graves. (2013). Speech recognition with deep recurrent neural networks, *In:* Acoustics, speech and signal processing (icassp), 2013 IEEE International Conference on, 2013, 6645—6649.

[81] Richard Socher, et al. (2013). Recursive deep models for semantic compositionality over a sentiment treebank, *In*: Proceedings of the conference on empirical methods in natural language processing (EMNLP), vol. 1631, 2013, p 1642.

[82] Richard., Lin., Cliff, C., Manning., Chri., Ng., Socher, Andrew Y. (2011). Parsing natural scenes and natural language with recursive neural networks, *In:* Proceedings of the 28th international conference on machine learning (ICML-11), 2011, 129 -136.

[83] Richard., Huval., Brody., Manning., Christopher, D., Ng., Socher, Andrew Y. (2012). Semantic compositionality through recursive matrix-vector spaces, *In*: Proceedings of the 2012 joint conference on empirical methods in natural language processing and computational natural language learning, 2012, p. 1201—1211.

[84] Yann., Bottou, L. (1998). Gradient-based learning applied to document recognition, *In:* Proceedings of the IEEE, p. 2278—2324.

[85] Yoon, Kim. (2014). Convolutional neural networks for sentence classification, *arXiv preprint arXiv:1408.5882*.

[86] Tomas., Chen., Kai., Corrado., Greg., Dean., Jeffrey Mikolov. (2013). Efficient estimation of word representations in vector space, *arXiv preprint arXiv:1301.3781*, 2013.

[87] Ruder, Sebastian, Ghaffari, Parsa, Breslin, John G. (2016). INSIGHT-1 at SemEval-2016 Task 5: Deep Learning for Multilingual Aspect-based Sentiment Analysis, *arXiv preprint arXiv:1609.02748*.

[88] Tai, Kai Sheng., Socher, Richard., Manning, Christopher D. (2015). Improved semantic representations from tree-structured long short-term memory networks, *arXiv preprint arXiv:1503.00075*, 2015.

[89] Al Sallab, Ahmad A., et al. (2015). Deep learning models for sentiment analysis in Arabic, in *ANLP Workshop 2015*, p 9.

[90] Borth,Damian., Ji, Rongrong., Chen, Tao., Breuel, Thomas., Chang, Shih-Fu. (2013). Large-scale visual sentiment ontology and detectors using adjective noun pairs," in *MM '13 Proceedings of the 21st ACM international conference on Multimedia*, Barcelona, Spain, Pages 223-232.

[91] Xu, Can., Cetintas, Suleyman., Lee, Kuang-Chih., Li, Li-Jia. (2014). Visual sentiment prediction with deep convolutional neural networks, *arXiv preprint arXiv:1411.5731*.

[92] You, Quanzeng., Luo, Jiebo, Jin, Hailin., Yang, Jianchao. (2015). Robust image sentiment analysis using progressively trained and domain transferred deep networks," *arXiv preprint arXiv:1509.06041*.

[93] Morency, Louis-Philippe., Mihalcea, Rada., Doshi, Payal. (2011). Towards multimodal sentiment analysis: Harvesting opinions from the web, In: *Proceedings of the 13th international conference on multimodal interfaces*, 2011, p 169-176.

[94] Poria, Soujanya., Cambria, Erik., Howard, Newton., Huang, Guang-Bin., Hussain, Amir. (2016). Fusing audio, visual and textual clues for sentiment analysis from multimodal content, *Neurocomputing*, volume 174, 50-59.

[95] Refaee, Eshrag., Rieser, Verena. (2014). An Arabic Twitter Corpus for Subjectivity and Sentiment Analysis., in *LREC*, 2014 p 2268-2273.

[96] Kaseb, Gehad S., Ahmed, Mona F. Arabic Sentiment Analysis approaches: An analytical survey.