



# Sentiment analysis in Arabic: A review of the literature



Naaïma Boudad\*, Rdouan Faizi, Rachid Oulad Haj Thami, Raddouane Chiheb

ENSIAS, Mohammed V University, Rabat, Morocco

## ARTICLE INFO

### Article history:

Received 22 January 2017

Revised 21 March 2017

Accepted 23 April 2017

Available online 21 July 2017

### Keywords:

Sentiment analysis

Opinion mining

Arabic

Machine learning

Lexicon based approach

## ABSTRACT

Within the last couple of years, Sentiment Analysis in Arabic has gained a considerable interest from the research community. In this respect, the objective of this paper is to provide a review of the major works that have dealt with this research area in this language. A thorough investigation of the available literature revealed that the works were mainly concentrated on dealing with specific Sentiment Analysis tasks. To this end, they used three different approaches, namely supervised, unsupervised and hybrid. The results that these studies achieved are interesting but divergent. This divergence is relatively due to the type of approach opted for, the task that is being analysed as well as to the specificities and intricacies of the Arabic variety understudy.

© 2017 Ain Shams University. Production and hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

Sentiment Analysis (also called Opinion Mining) refers to the use of Natural Language Processing and Machine Learning to identify and extract subjective information in a piece of writing. Sentiment Analysis is extremely useful as it enables us to gain an overview of the wider public opinions or attitudes towards certain topics, products or services.

Today, with the proliferation of reviews, ratings, recommendations and other forms of expression thanks to the emergence of the participative Web, Sentiment Analysis has become one of the essential research fields whose application is clearly visible in a variety of domains such as politics, commerce, tourism, education and health.

Given the importance of Sentiment Analysis, many research works have been devoted to this research area. However, most of these studies have focused on English and other Indo-European languages. Very few studies have, actually, addressed Sentiment Analysis in morphologically rich languages such as Arabic. Never-

theless, given the increasing number of Arabic internet users and the exponential growth of Arabic online content, Sentiment Analysis in this language has gained the attention of many researchers in the last decade.

The objective of this paper is, therefore, to review the major studies that have been conducted on Sentiment Analysis in Arabic. The rest of this paper is organized as follows. Section 2 gives background information about Sentiment Analysis. Section 3 outlines the linguistic characteristics of Standard Arabic. Section 4 describes the challenges of Sentiment Analysis in Arabic. Section 5 reviews the most important research studies carried out on Sentiment Analysis in Arabic. In Section 6, the findings are discussed and we conclude in Section 7.

## 2. Sentiment analysis: Background information

Sentiment Analysis is a research field that analyzes people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions towards entities such as products, services, organizations, individuals, issues, events, topics, and their attributes [1].

In Sentiment Analysis [1], an opinion is defined by making reference to a quintuple (**o**; **a**; **so**; **h**; **t**) that consists of:

- Object '**o**', which is the opinion target. It can be a product, service, topic, issue, person, organization, or an event.
- Aspect '**a**', which is the targeted attribute of the object '**o**'.
- Sentiment orientation '**so**' that indicates whether an opinion is positive, negative or neutral.

\* Corresponding author.

E-mail addresses: [naaïma.boudad@gmail.com](mailto:naaïma.boudad@gmail.com) (N. Boudad), [rdfaizi@gmail.com](mailto:rdfaizi@gmail.com) (R. Faizi), [rachid.ouladhajthami@gmail.com](mailto:rachid.ouladhajthami@gmail.com) (R. Oulad Haj Thami), [radchiheb@gmail.com](mailto:radchiheb@gmail.com) (R. Chiheb).

Peer review under responsibility of Ain Shams University.



Production and hosting by Elsevier

- Opinion holder '**h**', which is the person or organization that expresses an opinion.
- Time '**t**': the moment in which this opinion is expressed.

Sentiment Analysis is a highly challenging research area that involves different complex tasks. The most studied tasks are subjectivity classification, sentiment classification, lexicon creation, aspects extraction, aspect sentiment classification and opinion spam detection.

### 2.1. Levels of sentiment analysis

Sentiment Analysis can be investigated at three levels of granularity, namely document level, sentence level, and aspect level.

At the document level, the whole piece of writing is dealt with as one unit and is assigned to a positive, negative or neutral class. This level of analysis supposes that each document expresses an opinion on a single entity and has only one opinion holder. Therefore, it cannot be applied to documents that evaluate or compare multiple entities.

At the sentence level, Sentiment Analysis aims to identify whether the sentence holds an opinion or not, and to evaluate the sentiment orientation of subjective sentences. This level of granularity is more challenging by the fact that the sentiment orientation of words is highly context-dependent. Sentence level sentiment classification also deals with comparative and sarcastic sentences.

Aspect level performs finer-grained analysis. Its purpose is to discover the quintuple (**o**; **a**; **so**; **h**; **t**) of the opinion. The two key tasks in this level are aspect extraction and aspect sentiment classification. In aspect extraction, the attributes are identified. In aspect sentiment classification, the sentiment orientation of different aspects are defined. A positive opinion on an object can be positive on just an attribute of the object, but not on the object as a whole.

### 2.2. Approaches

Sentiment Analysis is usually carried out by using three typical approaches: supervised, unsupervised and hybrid.

The supervised approach is based on machine learning algorithms such as Support Vector Machine (SVM), Naïve Bayes (NB), Artificial Neural Networks (ANN) and K-Nearest Neighbour (KNN). This approach requires a large labelled data to train a classifier or a set of classifiers. The training process aims to build a classification model capable of predicting the sentiment of new texts. However, labelled data are often unavailable or a lot of efforts are required to annotate the collected data.

By contrast, the unsupervised approach makes use of the words and their corresponding sentiments. Each word in the lexicon has a score that represents its polarity, namely positive, negative or neutral. The lexicon could be created from existing dictionaries or from a corpus. Unlike the supervised approach, the unsupervised model does not require labelled datasets but it needs a wide-coverage lexicon that covers a maximum number of sentiment words.

Rather than using one of the above-mentioned approaches, some research studies have opted for a combination of both approaches. The new model, termed semi-supervised or weakly supervised approach, uses a great amount of unlabelled data and partially annotated data to build better classifiers. The system classifies unlabelled data using a supervised classifier trained with labelled data.

## 3. Arabic language

Given that Sentiment Analysis depends greatly on the morphology of the language being analysed, our objective in the present

section is to give a brief description of Arabic and to detail the linguistic features that make this language one of the most challenging varieties for Sentiment Analysis researchers.

Arabic is one of the six official languages of the United Nations. It is the official language of 27 countries and is spoken by more than 422 million people in the Arab world. On the web, Arabic is ranked the fourth mostly used language and the fastest growing during the last five years with a growth rate of 6091.9% in the number of Internet users.

Arabic has three main varieties: Classical Arabic; which is the language of the Qur'an (Islam's Holy Book); Modern Standard Arabic (MSA) and dialectal Arabic. MSA the most eloquent Arabic language variety used in writing and in most formal speech. Dialectal or colloquial Arabic refers to all oral varieties spoken in daily communication. These vary from one Arab country to another and from one region of the same country to another.

### 3.1. Arabic orthography

Unlike Latin languages, Arabic is written from right to left and is distinguished by the absence upper or lower cases. Its alphabet includes 28 letters: 25 consonants and only 3 vowels. But in addition to these vocal segments, the Arabic script uses diacritical marks as short vowels. These are placed either above or below the letters to provide the correct pronunciation and clarify the meaning of the word. The majority of MSA texts are written without short vowels. This is so because proficient speakers do not need diacritical marks in order to understand a given text. However, diacritical marks are often used in children's books as well as books for Arabic learners. The absence of diacritical marks in the majority of texts presents a lexical ambiguity problem that challenges computational systems. For example the undiacritized word شعر may mean (شِعْرٌ poetry), (شَعْرٌ hair) or (شَعُرٌ to feel).

### 3.2. Arabic morphology

Arabic language has a very complex and rich morphology in which a word may carry important information. As a space delimited token, a word in Arabic reveals several morphological aspects: derivation, inflection, and agglutination.

#### 3.2.1. Derivational morphology

Derivational morphology is the mechanism of creating a new word based on an existing word with a part of speech possibly different, e.g. in English the adjective "weekly" is derived from the noun "week".

Like other Semitic languages, Arabic Morphology consists of a root-and-pattern representation. All Arabic words are based on a "root", which is a sequence of consonants that contain the base meaning of the word. The vowels and non-root consonants are added following specific patterns to create a variety of related words. For example, the three letters "ktb" is a root that means, "write". If it is put in the pattern "1a2a3a" (kataba, كَتَبَ) where the numbers correspond to the root letters. By adding the long vowel ("a:" after the first letter we get a new pattern "1a:2a3a" and a new verb (ka:taba, كَاتَبَ) which means "correspond". Table 1 lists some words derived from the root "ktb" together with their meanings.

#### 3.2.2. Inflectional morphology

Inflectional morphology defines the variation of a word to describe the same meaning in different grammatical categories (e.g. in English: write, wrote, written). The set of these inflected word-forms is called a lexeme class. To represent the lexeme, a lemma, which is a particular form, is conventionally selected.

**Table 1**  
Words derived from the root “ktb”.

Words derived from the root		Pattern	Meaning
كُتِبَ	kataba	1a2a3a	Write
كَاتَبَ	ka:taba	1a:2a3a	Correspond
مَكْتَب	maktab	Ma12a3	Desk
كُتُب	Kutub	1u2u3	Books
كَاتِب	ka:tib	1a:2i3	Writer

In Arabic, words inflect for seven categories: tense (past and present), person (1st, 2nd and 3rd), number (singular, dual and plural), gender (feminine and masculine), case (nominative, accusative and genitive), mood (indicative, imperative, subjunctive, jussive and energetic), and voice (active and passive). Table 2 presents the inflection of the verb “ktb” (write) depending on tense, person, number and gender.

### 3.2.3. Agglutinative morphology

Arabic is an agglutinative language, which means that the word may be attached a set of clitics (affixes). These clitics are divided into 4 classes (cf. Table 3) and apply to a word base in a strict order:

CONJ + PART + DET + BASE + PRON

The English phrase “and with his work”, for example, corresponds to the Arabic form “ويعمله”. This word can be split into four parts (و + ب + عمل + ه): The conjunction “و” “and”, the particle proclitic “ب” “with”, the stem or the word base “عمل” “work”, and the possessive pronoun “ه” “his”. Multiple word prefixing, suffixing and affixing generate different word from the same stem.

The complexity of the Arabic word structure is one of the main difficulties that researchers face when dealing with Arabic Sentiment Analysis. The next section examines some of the key challenges facing efforts to set up an accurate system for Arabic.

## 4. Sentiment analysis in Arabic: Challenges

### 4.1. Morphological analysis

Morphological analysis is an important phase in Sentiment analysis. Its main purpose is to decompose words into morphemes and to associate each morpheme with a morphological information such as stem, root, POS (Part Of Speech), and affix. As we saw in the previous section, Arabic is a morphologically complex language. This complexity requires the development of appropriate systems that are able to deal with tokenization, spell checking, stemming, lemmatization, pattern matching, and part-of-speech tagging. Nowadays, many morphological analysers for Arabic are already developed; some of these are freely available while the rest have a commercial purpose. Among those referred to the literature are Xerox Arabic Morphological Analysis and Generation [2], and Buckwalter Arabic Morphological Analyser (BAMA) [3]. However, these

**Table 2**  
Inflection of the verb ‘ktb’ (write).

	Past		Present	
1st person Singular	Katabtu	كُتِبْتُ	Aktubu	أَكْتُبُ
1st person Plural	Katabna:	كُتِبْنَا	Naktubu	نَكْتُبُ
2nd person, Masculine Singular	Katabta	كَتَبْتَ	Taktubu	تَكْتُبُ
2nd person Feminine Singular	Katabti	كَتَبْتِ	Taktubin	تَكْتُبِينَ
2nd person Dual	Kabtuma:	كُتِبْتُمَا	Taktuba:n	تَكْتُبَانِ
2nd person Masculine Plural	Kabtum	كَتَبْتُمْ	Taktubu:n	تَكْتُبُونَ
2nd person, Feminine Plural	Kabtunna	كَتَبْتُنَّ	Taktubunna	تَكْتُبْنَ

**Table 3**  
Classes of Arabic clitics.

Conjunction proclitics	+“” و +“” ف	“w” “f”	And Then
Particle proclitics	+“” ال +“” ب +“” ك +“” س	“l” “b” “k” “s”	to/for by/with as/such will/future
Definite article	+“” ال	“al”	The
Pronominal enclitics	+“” ه +“” ها +“” هم +“” هما +“” هن +“” ك +“” كم +“” كما +“” كن +“” نا +“” ي	“h” “ha:” “hum” “huma:” “hunna” “k” “kum” “kuma:” “kunna” “na:” “y”	His, its, him, it, Her, its, it, him Their, Them (Male, Plural more than 2) Their, Them (Double) Their, Them (Female, Plural more than 2) Your, you (single) Your, you (Male, Plural more than 2) Your, you (Double) Your, you (Female, Plural more than 2) Our, us My, me

systems suffer from significant limitations especially in handling ambiguity that can result from the omission of diacritics (vowels), the free word-order nature of Arabic sentence, or the presence of an elliptic personal pronoun.

### 4.2. Dialectal Arabic

For communication purposes, Arabic speakers usually use colloquial Arabic rather than MSA. There are around 30 major Arabic dialects that differ from MSA and from each other phonologically, morphologically, and lexically [4]. Furthermore, Arabic dialects have no standard orthographies and no language academies. Therefore, using tools and resources designed for MSA to process Arabic dialects generates considerably low performance. Recently, researchers have started developing parsers for specific dialects such as CALIMA [5] for Egyptian Dialect. However, these analysers still have low accuracy and are made only for particular dialects. Filling this gap in processing Arabic will improve information retrieval effectiveness specifically for social media data.

### 4.3. Arabizi

Arabizi, Arabish, or Romanized Arabic refers to a system of writing Arabic using Latin characters. It is widely used to write MSA as well as Arabic dialects in social media platforms. Dealing with this form of writing has been the subject solely of studies that aim at detecting and converting Arabizi into Arabic.

As far as Sentiment Analysis is concerned, the published works have not dealt with this problem, as texts are pre-processed to filter out all Latin letters. To the best of our knowledge, [6] is the only published work that has handled this task.

#### 4.4. Named entity recognition

In Arabic, large portions of Arabic names are associated with positive adjectives. For example, the first name "سعيد" corresponds to the adjective "سعيد" which means "happy". In addition, Arabic proper nouns are not capitalized as in Latin languages, a fact which complicates the identification of named entities. For this reason, a system of Named Entity Recognition is crucial in analysing Arabic texts and distinguishing between entity names and sentiment words.

### 5. Literature review

The objective of the present section is to review the most important research studies that have dealt with Sentiment Analysis in Arabic. Note, however, that the works that we shed light on below are not classified on the basis of the approaches or techniques that they opted for in their analyses but on the specific Sentiment Analysis task or tasks that every work is concerned with.

#### 5.1. Subjectivity classification

Subjectivity Classification is the task of determining whether a text is subjective or objective. Such a classification of texts can be performed using machine learning as well as lexicon-based approaches.

In this respect, Abdul-Mageed et al. [7,8] proposed a machine learning approach to perform a subjectivity and sentiment classification at the sentence level. The first work dealt with a dataset of newswire documents extracted from PATB and manually annotated (1281 OBJ and 1574 SUBJ). In [8], the authors collected and annotated 11,918 sentences from different types of social media services. Classification is effected in two stages. In the first stage, a distinction is drawn between a subjective and an objective text. (i.e. Subjectivity Classification). In the second stage, a difference is made between a positive and negative sentiment (i.e., Sentiment Classification). In this work, Abdul-Mageed et al. used SVM as a learning algorithm together with language specific and general features. Language-independent features include n-grams, domain, unique and polarity lexicon features. Arabic specific features were added to investigate the impact of morphological information on the performance. The findings indicated that using POS tagging and lemmas or lexemes to extract the base forms of words has a positive impact on subjectivity and sentiment classification.

Nabil et al. [9] presented a 4 way sentiment classification that classifies texts in four classes: objective, subjective negative, subjective positive and subjective mixed. Their dataset has 10,006 Arabic Tweets manually annotated using Amazon Mechanical Turk (AMT) service. They applied a wide range of machine learning algorithms (SVM, MBN, BNB, KNN, stochastic gradient descent) on the balanced and unbalanced datasets. However, using n-grams as unique features in multi-way classification did not give good results.

A lexicon-based approach was proposed in [10] to perform subjectivity classification of both MSA news articles and dialectal Arabic microblogs from Twitter. In order to build a large lexicon, the authors use two available lexicons: MPQA which is an existing English subjectivity lexicon, and ArSenti, a manually created Arabic lexicon. The first one is translated into Arabic using Machine Translation, and the second is automatically extended using a random graph walk method. All the words in Tweets and in the lexicon were tokenized and stemmed. Polarity stems, as indicated in the lexicon, were used as input feature vector to the learning module.

#### 5.2. Sentiment classification

Sentiment Classification aims at classifying subjective texts in two or more categories. The binary classification determines whether the text expresses a positive or a negative opinion. A multi-way classification labels texts according to the strength of the expressed sentiment, namely extremely negative, negative, neutral, positive or extremely positive.

Mountassir et al. [11] conducted a binary sentiment classification using three classifiers: NB, SVM and KNN. Two corpora were used: the first is developed by these authors and is composed of two domain-specific datasets (movies and sports). The second is OCA, a corpus of movie reviews developed by Rushdi-Saleh et al. [12]. Before the classification phase, the authors performed a pre-processing task by removing stop words, separating words from their clitics, eliminating terms used only once or twice in the dataset, and by replacing words by their stems. The authors found out that pre-processing, n-grams combination, and presence-based weighting improve the classification performance.

In [19], two forms of Sentiment Classification were explored: the polarity classification, which classifies reviews as either having a positive or negative sentiment, and the rating classification whose objective is to predict the rating of the review on a scale of 1–5. Aly and Atiya [19] created a Large-scale Arabic Book Review (LABR), a dataset of over 63,257 book reviews collected from [www.goodreads.com](http://www.goodreads.com). Reviews with rating 4 or 5 were labelled as positive. Negative reviews were those with rating 1 or 2 while reviews that were rated 3 were considered neutral. Since the number of positive reviews (42,832) was much larger than that of negative reviews (8224), they applied machine learning in both balanced and unbalanced data using SVM, MNB and BNB as algorithms and n-grams as features. For sentiment polarity classification, the evaluation of their dataset achieved quite good results (~90% accuracy), but for rating classification there is much room for improvement (~50% accuracy).

For their parts, El-Baltagy et al. [25] proposed a lexicon-based approach to establish a sentiment classification of Egyptian Arabic texts. After building a lexicon of 4392 terms, the authors used two datasets (Twitter dataset of 500 tweets and Dostour dataset of 100 web comments) to evaluate two unsupervised classification algorithms. The first calculates one score for each document by adding up weights of negative and positive terms. The second algorithm assigns a positive and a negative weight to each term in the lexicon and calculates positive and negative scores for each document. The authors achieved good results using the two algorithms on a Twitter dataset (83.8% accuracy).

In [48], El-Makky et al. combined Sentiment Orientation (SO) algorithms with a machine learning classifier to propose a hybrid approach. For each document in a Twitter dataset, they used the lexicon-based approach to compute Sentiment Orientation scores. These scores were integrated with different features such as uni-grams, language independent features, Tweets-specific features and stem polarity features so as to create an input feature vector for the SVM classifier. This combination of the Machine Learning classification approach and the lexicon based approach led to slightly better results than a one-approach result (accuracy 84%).

Determining the sentiment intensity of Arabic phrases was the goal of task 7 of SemEval 2016 [49]. The task aims to provide a score between 0 and 1 that indicates the Sentiment Intensity (SI) of a phrase. A score of 1 indicates the maximum of "positive strength" and a score of 0 indicates the maximum of "negative strength". The **SemEval organizers** provided the participants with a development set of 200 terms commonly found in Arabic tweets and a set of 1166 terms for the evaluation period. Three systems were submitted from three teams: NileTMRG [50], iLab-Edinburgh [51] and LSIS [52].



Using a supervised approach, the NileTMRG team [50] collected 249 K tweets by querying Twitter using test set terms. Then, they classified this collection using their own sentiment analyzer [42] developed with Complement Naïve Bayes classifier [53] and trained on 11,242 Arabic tweets. To assign the Sentiment Intensity to each term, the normalized point wise mutual information relative the positive class was computed and re-scaled so that the values would range from 0 to 1.

The system proposed by the iLab-Edinburgh team [51] employs a hybrid approach of rule-based methods and supervised learning. The first phase of the system uses Linear Regression trained with the labMT1.0 Sentiment Lexicon [54], which is a publicly available list of 10k Arabic positive/negative words. Each entry of the Lexicon is associated to its Sentiment Intensity score ranged between 1 and 9 (1 very-negative/9 very-positive). After re-scaling the Sentiment Intensity score to [0, 1], the Linear Regression model is used to calculate an initial Sentiment Intensity score. In the rules-based phase of the system, a set of hand-crafted rules and a combination of three publicly available sentiment lexica, namely ArabSenti [55], MPQA [56], and Dialect lexical [57] was used to adjust initial SI scores.

The third team [52] proposed an unsupervised method. Their system computes the degree of dependence between a term and a positive class in sentiment lexica using the pointwise mutual information. When the term is not found in the lexica, a web search engine was used to calculate its sentiment orientation based on its co-occurrence near positive or negative words.

Evaluation showed that the iLab-Edinburgh team presented the best performing system and the results obtained using supervised methods are distinctly higher than the results got using unsupervised methods. Nevertheless, the results achieved on Arabic Twitter data was markedly lower than the results reached on a similar English Twitter data set [49].

### 5.3. Aspect based sentiment analysis

Compared to the above mentioned tasks, Aspect Based Sentiment Analysis (ABSA) is the least studied in Sentiment Analysis in spite of its crucial importance. This challenging task aims to recognize entities (e.g. brand of a mobile phone) and their aspects (e.g., 'battery', 'screen') and to estimate the sentiment expressed towards each aspect. As far as the Arabic language is concerned, only very few research papers have been published on ABSA.

One of the earliest research works that dealt with this task is [58]. In this work, Alhazmi and Salim introduced a supervised approach to extract the opinion target from Arabic Tweets. To build a training dataset, they manually tagged the opinion target in 500 collected Tweets. After pre-processing Tweets, each word was considered as a training vector defined by POS, named entities, English words and tweet hash tags features. Classification was carried out by specifying that a given word is either an opinion target or not. Experiments were undertaken using three classifiers: Naïve Bayes, Support Vector Machine and K-Nearest Neighbour. The best result was reached using the K-Nearest Neighbour classifier with an F-Measure of 91%.

Within the same task, Al-Smadi et al. [59] set up a benchmark dataset of Arabic reviews. Their Human Annotated Arabic Dataset (HAAD) contains 1513 reviews selected from LABR [19] and manually annotated. Following SemEval2014 guidelines proposed in [60], annotators were asked, in the first phase, to identify aspect terms and provide the polarity for each aspect term. In the second phase, annotators recognized aspect categories and their polarities. As a benchmark dataset, the authors conducted an evaluation baseline of four tasks: aspect term extraction, aspect term polarity, aspect category extraction, and aspect category polarity. The adopted approach uses a majority baseline that assigns the most

repeated polarity in the training data to all aspect terms and categories.

Al-Smadi et al.'s work was improved by Obaidat et al. [61] using a Lexicon-Based Approach to deal with aspect category extraction and aspect category polarity. The proposed approach for aspect category polarity employed a polarity lexicon extracted from the training set. The lexicon assigns every word in the training set to the class corresponding to its highest occurrence frequency. For word missing in the training set, the authors made use of machine translation and POS tagger to get polarity from SentiWordNet. The accuracy rate obtained with this approach (i.e. 71%) outperforms significantly the baseline results (42.6%). To deal with aspect category extraction, the authors created a category lexicon using seed words and Pointwise Mutual Information (PMI). The trained words were assigned to the category with the most similar seed words. Each entry in the lexicon was weighted by its PMI value. The category of a test review was defined by summing the weights of its words and selecting the aspect category with the highest sum. This approach showed an improvement of 8.2% over the baseline results.

### 5.4. Building resources

With the rising interest Sentiment Analysis in Arabic, developing multi-genre corpora and large Arabic Lexica for word-level sentiment evaluation has become of utmost importance. In this respect, several research studies have recently made attempts to make large Arabic resources available for Sentiment Analysis researchers.

Abdul-Mageed and Diab [55], for instance, presented a large scale multi-genre sentiment lexicon. This lexicon is made up of 224,564 entries covering MSA and multiple Arabic dialects. The authors collected and manually tagged two word lists from both Penn Arabic Treebank and Yahoo Maktoob. Lists were automatically developed using Google's translation API of three existing English lexica: SentiWordNet, YouTube Lexicon, and General Inquirer. To expand the lexicon's coverage, they used a statistical method based on PMI to extract other polarized tokens from both Twitter and chat datasets. Despite the large size of the resulting resource, many of the entries are neither lemmatized nor diacritized, which limits the usability of their lexicon.

In their attempt to build Arabic multi-domain resources for Sentiment Analysis, ElSahar and El-Beltagy [37] proposed a semi-supervised approach to generate multi-domain lexica out of four multi-domains reviews datasets. This method makes use of the feature selection capabilities of SVM to select the most efficient unigram and bigram features. Although the created lexicon covers a variety of domains, it was extracted only from reviews, which restricts its usefulness just for social media Sentiment Analysis.

Badaro et al. [66] set up ArSenL, a lexicon for Arabic sentiments using two approaches based on English SentiWordNet (ESWN). The first method links each term in ArabicWordNet, on the one hand, with ESWN to get sentiment scores, and on the other hand with SAMA (Standard Arabic Morphological Analyser) to find the correct lemma forms. In the second approach, English glosses associated with SAMA's entry were explored automatically to find the most similar synset in ESWN. The union of the two resulting lexica has a good coverage but is limited to MSA.

In [62], AWATIF a multi-genre corpus of MSA was collected from three different sources: Penn Arabic Treebank (PATB), Wikipedia Talk Pages, and web forums. One part of the corpus was labelled using crowdsourcing on Amazon Mechanical Turk (AMT) annotation. Annotation of the second part was carried out by students who received specific guidelines. The last part was also labelled by students but with simple guidelines. Abdul-Mageed and Diab observed that the value of inter-annotator agreement

improves using specific guidelines. However, this corpus is not publicly available and is not used in any other work.

In addition to these corpora, an Arabic Twitter corpus was collected in [57] using Twitter API and cleaned in a pre-processing phase. Two native speakers of Arabic annotated manually 8868 Tweets using four labels: neutral, mixed, positive and negative. Morphological, syntactic, and semantic features were also added to the annotation.

### 5.5. Opinion holder extraction

An Opinion holder is the person or the organization that expresses an opinion. In most social media platforms, opinion holders are usually the authors of reviews. Therefore, we do not need to extract them. In some cases, as in news articles, the person or organization that holds an opinion is often explicitly stated in the actual text and needs to be extracted. Nevertheless, extracting opinion holders has drawn little attention from Arabic opinion mining researchers. To the best of our knowledge, there are only two research works that have been concerned with this issue, namely [73,74].

In [73], the authors explored the problem in Arabic news articles using three different approaches. The first approach is semi-supervised and uses a set of handcrafted patterns. POS tags and key phrases were used to define 43 patterns, which were chosen to run a pattern matcher code and identify opinion sources on the tested data. The second is a supervised machine learning approach that uses the Conditional Random Field classifier (CRF). To train this classifier, the authors used features such as surrounding words, POS, Named Entity and sentiment words. The third approach is a combination of the two previous approaches using patterns such as CRF features. Their experiments showed that the CRF outperforms patterns in terms of recall and precision. Moreover, adding patterns as a feature to CRF is insignificant compared to other features such as the Named Entity feature.

### 5.6. Opinion spam detection

Opinion spam refers to false reviews written to promote a low quality product using positive opinions or to damage the reputation of a given product with negative opinions. The objective of this task is to detect automatically spam reviews using techniques that usually depend on three types of features: content of the review, review meta-data, and real-life knowledge about the product.

With the growing popularity of online reviews or comments and their major impact on business decision making, this illegal activity has become more developed in terms of both quality and quantity. The number of spam opinions have actually increased and the reviews are getting more and more sophisticated. The latter fact, actually, justifies the scarcity of studies devoted to this issue in Arabic.

A preliminary study that was concerned with opinion spam detection was conducted by Wahsheh et al. in [75]. The study was based on 3090 Arabic opinions collected from Yahoo-Maktoob social network. The authors employed ACLWSDS, which is an Arabic spam URL detection system developed in [76]. Opinions containing a URL were classified, either as high-level spam if the URL was considered as spam by ACLWSDS, or as low-level spam if the URL was considered as non-spam. In the absence of URLs and some specific metrics, the opinion was categorized as a non-spam. Evaluating this method with SVM algorithm achieved favourable results. However, using only the URL filtering technique may not be efficient.

In [77], Abu Hammad and El-Halees proposed a supervised approach for detecting opinion spams in Arabic. This approach combines techniques from data and text mining. The authors col-

lected 2848 Arabic reviews from online accommodation booking websites namely, booking.com, tripadvisor.com, and agoda.ae. In addition, they integrated their dataset into a coherent form data and labelled each review with a spam or a non-spam label. For classification purposes, they used NB, K-NN and SVM classifiers with 10 folds cross-validation. Although their system was limited to hotel reviews, it was able to generate a high accuracy by combining data and text classification.

## 6. Discussion

Although research on Sentiment Analysis has started in Arabic since 2008, studies on this issue have witnessed a rapid increase in the last three years. This is mainly due to the exponential growth of Arabic text content on the web. The works that have been reviewed in the section above have revealed that researchers have dealt with different Sentiment Analysis tasks: Subjectivity Classification, Sentiment Classification, Aspect based Sentiment Analysis, building resources, extracting opinion holders, and detecting spam opinions.

Subjectivity Classification has been carried out in a very limited number of works. It is treated as an initial phase of sentiment classification to filter objective elements and not as a standalone issue.

Sentiment Classification is the most widely studied topic. As indicated in Table 4, Sentiment Classification at document level has been the subject of a lot of research studies. The limited number of studies on sentence level may due to the fact that it requires the detection of sentence boundaries which is another Arabic NLP problem. For both levels, the supervised learning approach outperformed unsupervised and semi-supervised methods. Note also that some supervised classifiers such as SVM and NB have repeatedly been applied. The effectiveness of these classifiers is probably the reason behind this choice. The same thing can be said about the used features; word stem and n-grams are frequently chosen as features. Regarding datasets, each researcher built and used his own dataset. Thus, no common dataset was used for benchmarking results and evaluating experiments. Early research analysed texts collected from the web and focused on MSA language. By contrast, researchers have recently treated principally social media texts and dealt with Arabic dialects as well as MSA.

Another issue that has attracted the attention of researchers concerns the building of Arabic lexica and corpora resources. Table 5 presents a summary of research studies related to this research area. In this respect, it is worthwhile to note that so as to compile lexicons, researchers have adopted two main methods:

- Corpus based approach: Extension of a seed list of sentiment words by identifying more sentiment words and their orientations from a corpus.
- Dictionary based approach: Extension of few sentiment words by searching in an available dictionary such as Arabic WordNet for their synonyms and antonyms.

Seed words can be collected manually or using Machine Translation of English lexica such as SentiWordnet. However, unavailability of dialectal Arabic dictionaries has restricted the use of the second approach to MSA.

Note also that the most of works introduced lexicons with word or terms entries and not with diacritized lemmas, which calls into question the effectiveness and the real size of the lexicon.

Very few attempts have, however, been concerned with Aspect based Sentiment Analysis, opinion holders' extraction and spam opinion detection in the Arabic language. These tasks still need much deeper investigation and research.

**Table 4**  
Summary of sentiment classification works.

Reference	Year	Approach	Level	Dataset	Algorithm	Features	Language	Accuracy	Advantages and disadvantages
[13]	2008	Supervised	Document	Web forum	Classification: SVML Feature selection: Genetic algorithm (GA), information-gain (IG)	Word Roots + n-gram, punctuations, Stylistic features	MSA	93.60%	+ Wide Feature Set + Using Feature selection technique POS Feature not used-limited to one domain
[14]	2010	Unsupervised	Sentence	Reviews	Based Lexicon	Lexicon scores	MSA	43%	+ Sentence level – Created lexicon include scoring garbage words
[12]	2011	Supervised	Document	Movie reviews	SVM, NB	Stem n-grams/word n-grams	MSA	90%	– Could try more features
[7]	2011	Supervised	Sentence	Penn Arabic Treebank (PATB)	SVM	Polarity Lexicon, n-gram, Surface, lemma and stem form of word, -Morphological Features: person, state, gender, tense, aspect, and number	MSA	95% (F-Mesure)	+ Using morphological Features – Feature selection technique is not used
[15]	2012	Unsupervised	Document	Twitter	Based Lexicon	Lexicon polarity	MSA + saudi dialect	73%.	+ Unsupervised approach: not need of training data – Small size of lexicon
[16]	2012	Supervised	Document	Aljazeera	SVM, BN, KNN	Word stem + n-grams	MSA	93% F1	– Could include more features
[11]	2012	Supervised	Document	Aljazeera	SVM, NB, KNN	Stems words	MSA	74.5%	+ Address the problem of unbalanced – Data sets COULD include more features
[17]	2012	Supervised	Sentence	Twitter	SVM, BN	n-gram	MSA + Egyptian dialect	72.6%	+ Dealing with Egyptian dialect – Could use more algorithms
[18]	2013	Supervised	Document	movie reviews + PATB	SVM	Word uni-gram, Roots uni-gram	MSA	93.2%	+ Evaluating different Rooting Libraries – Unbalanced dataset
[10]	2013	Supervised	Document	Twitter	NB	Word stems, POS, lexicon polarity, n-grams, tweet specific features	MSA, DA	80.60%	+Using POS feature – Negation not handled
[19]	2013	Supervised	Document	Book reviews	SVM, MBN, BNB	n-grams	MSA	91%	+ Large size of training dataset – Could use more features
[20]	2013	Supervised	Document	Twitter	SVM, NB	Word stem + n-grams	MSA	87.43%	+ Evaluating the add value of each feature – Small and unbalanced data sets
[21]	2013	Unsupervised	Document	Social network	Based Lexicon	Lexicon polarity	MSA, DA	66.2%	– Negation and Intensification are not considered + Multi domain lexicon
[22]	2013	Supervised	Document	Facebook	SVM, NB, Decision Tree	Common Words between Post and Comments	MSA	73.4%	+ Exploring new features – Could use more classifiers
[23]	2013	Supervised	Document	Facebook	SVM, NB	Combining n-grams	MSA	75.31%	+ Using PoS Tagging – Small data set
[24]	2013	Supervised	Document	Product reviews	SVM, NB, Rocchio, (voting, Stacking classifiers)	Word stem	MSA, DA	97.53% (F macro)	+ Combining classifiers – Could use more features

Table 4 (continued)

Reference	Year	Approach	Level	Dataset	Algorithm	Features	Language	Accuracy	Advantages and disadvantages
[25]	2013	Unsupervised	Document	Twitter	Based Lexicon	Sentiment orientation weight	MSA, Egyptian dialect	83%	+ Calculating Sentiment orientation – Negation and Intensification are not considered
[26]	2014	Supervised	Document	Social network	Based Lexicon	Sentiment orientation weight	MSA, DA	97%	+ Dealing with polarity and Intensity of sentiment – Extracting and weighting features was done manually
[27]	2014	Supervised	Document	Aljazeera	SVM, NB, KNN	Word stem + n-grams	MSA, DA	96%	+ Using features correlation – Could use more features
[28]	2014	Supervised	Document	Twitter	SVM, NB, KNN	Word stem + n-grams	MSA, DA	67%	– Could use more features and more classifiers
[29]	2014	Supervised	Document	Facebook and blogs	SVM, NB	Word stem + n-grams	MSA, DA	75%	– Could use more features and more classifiers
[9]	2015	Supervised	Document	Twitter	SVM, MNB, BNB, KNN, passive aggression, SGD, logistic regression,	n-grams	MSA, DA	69.10%	+ Dealing with 4 way classification – No preprocessing
[30]	2015	Unsupervised	Document	Twitter	Based Lexicon	Semantic sentiment weight + word stem	MSA, DA	46%	+ Building polarity lexicon – Negation and Intensification are not considered
[31]	2015	Hybrid	Document	Twitter	Based Lexicon + SVM	n-grams	MSA, DA	84.01%	+ Hybrid approach – Negation and Intensification are not considered
[32]	2015	Supervised	Document	Twitter	SVM, NB	Word stem	MSA, DA	87% (F-Mesure)	+ Using lexicon dialect to translate dialect word to MSA – Could use more features and more classifiers
[33]	2015	Supervised	Document	goodreads.com	J48, DecisionTable, KNN, SVM, NB, MNB	Word stem	MSA	58%	+ Dealing with Multi-way Classification – Could use more features
[34]	2015	Supervised	Document	Aljazeera + movie reviews	SVM, MaxEnt and ANN	Opinion features, discourse markers, stylistic features, domain dependent features and morpho-lexical features	MSA	85.06%	+ Combining classifiers – PoS feature is not used
[35]	2015	Supervised	Document	YouTube comments	NB, SVM, DT, based lexicon	Unigram, bigram, word stem, Sentiment orientation weight	MSA	95%	+ Fusing feature selection and lexicon
[36]	2015	Unsupervised	Document	Twitter	Machine Translation, Stanford Sentiment Classifier	NA	MSA	71%	+ Evaluating the effectiveness of Machine Translation – Limited to one English Sentiment Classifier
[37]	2015	Semi-supervised	Document	Multi domain reviews	Linear SVM, Logistic regression, Bernoulli Naive Bayes, KNN and stochastic gradient descent.	n-grams, Lexicon entries	MSA, DA	88%	+ Large multi-domain datasets – No preprocessing
[38]	2015	Supervised	Document	Hotel reviews	SVM	Low level stem	MSA	83%	+ New mathematical approach to compute feature weights – Limited to one feature
[39]	2015	Supervised	Document	Tweets, reviews, comment	SVM	Standard features, sentence-level features, linguistic features,	MSA, DA	95%	+ Extensive set of features – Feature selection not used



**Table 4** (continued)

Reference	Year	Approach	Level	Dataset	Algorithm	Features	Language	Accuracy	Advantages and disadvantages
[40]	2016	Supervised	Document	Book reviews	Hierarchical Classifieraction using: SVM, KNN, DT, NB	Bag-Of-Word	MSA	57.8%	+ Dealing with Hierarchical Classifieraction – Limited to Bag-Of-Word as feature
[41]	2016	Unsupervised	Document	Multi Domain reviews	Based-lexicon	SentiStrength	MSA-DA	62%	+ Using existing English tools – No add value
[42]	2016	Supervised	Document	Twitter	Complement Naïve Bayes	Sentiment Lexicon features, emoticon based features, text related features	MSA-DA	79.4%	+ Extensive set of features – PoS feature is not used
[43]	2016	Supervised	Document	Twitter	SVM, NB	n-gram	MSA-Saudi Dialect	89%	– Could use more classifiers and more feautres
[44]	2016	Supervised	Document	Hotels reviews, Facebook, Twitter, and YouTube.	SVM, Back-Propagation Neural Networks (BPNN), NB, Decision Tree	Word stem	MSA	96.06%	+ Using POS tagging – Could use more features
[45]	2016	Supervised	Document	Arabic tweets and Facebook comments	SVM, KNN, BN	Word stem + n-grams	Arabic	69.97%	– Could use more classifiers and more features
[46]	2016	Supervised	Document	Twitter	NB and DT	Word stem	MSA	64.85%	+ Spelling correction and removing proper nouns are used at pre-processing step – Could use more features
[47]	2016	Unsupervised	Aspect level	Reviews (from Forums, Facebook, YouTube)	Domain ontology, lexicon based, rules based	Part Of Speech, feature and sentiment weights	MSA	92.15%	+ Handling Aspect level + New methods: ontology, rules based

**Table 5**

Major works that built resources.

Reference	Type of resource	Size	Method	Data source used	Language
[62]	Corpora	5382 sentences	Manual annotation + crowd-sourcing	PATB + wiki + web forums	MSA
[63]	Lexicon	3479 words	Manual extension of 300 seed word	SentiStrength	MSA
[64]	Lexicon	8000 terms	Based corpus Method	Reviews and comments	MSA + Multi-dialects
[65]	Lexicon	16,800 words	– Machine Translation of SentiStrength – Based corpus Method	Twitter + Maktoob	MSA + Egyptian, khaliji, Levantine
[55]	Lexicon	224,564 terms	Machine Translation of English Lexicons + PMI method on collected corpus	Twitter + Maktoob	MSA + Egyptian and Levantine
[66]	Lexicon	28,780 lemma	Linking AWN and SAMA to ESWN	ESWN	MSA
[67]	Lexicon	8133 lemma	Translation and Extension of MPQA lexicon	Movie reviews	MSA
[68]	Lexicon	7576 words	Extension of 14 seed word using Arabic WordNet	Movies reviews	MSA
[69]	Idioms-Proverbs	3632 phrases	Collecting proverbs, Manual selecting end annotating commonly used phrases	Twitter + Microblogs + web forum	MSA + Egyptian dialect
[70]	Lexicon				
[70]	Corpora	4700 sentences	Twitter API and manual annotation	Twitter	MSA + Saudi dialect
[71]	Lexicon	34,821 entries	Linking a morphological analyzer (AraMorph) to ESWN	ESWN	MSA
[72]	Lexicon	120,000 Arabic terms	Crawling news websites, stemming, machine translation and querying English sentiment	Twitter	MSA
[25]	Lexicon	4392 terms	Extension of 380 word using	Twitter	MSA + Egyptian dialect
[37]	Domain specific lexicons	1913 entries	Semi-supervised approach using 1-norm Support Vector Machines	Multi domain Reviews	MSA + multidialect

The present study clearly shows that, for Arabic Sentiment Analysis, the use of machine learning techniques dominated other Lexicon based and hybrid techniques. In their survey, Medhat et al. [78] have found that the first works of English Sentiment Analysis used ML approaches but in the last few years, researchers are heading towards the lexicon based approach. Supervised algorithms yielded the best performance in most of the studies. This performance depends heavily on the choice of algorithms, the quantity of training data and feature selection.

On binary classification, the results of most of the studies [26–28,33,34,41–45] demonstrated that Naïve Bayes and SVM algorithms performed better than other algorithms. In [43], SVM achieved the highest performances: accuracy (96.06%), precision (95.80%) and recall (96.40%), followed by NB with: accuracy (88.38%), precision (92.62%) and recall (84.99%). These results showed that SVM is the best adapted for large-scale and unbalanced data. However, NB performed better than SVM when feature selection techniques are used. Reducing feature correlation improves independence between the features and then enhances Naïve Bayes performance. Duwairi et al. [26] found that using Pearson coefficient for feature reduction makes Naïve Bayes accuracy increases from 53.80% to 78.64%, and SVM accuracy increases from 59.82% to 73.90%.

On Multi-Way classification, the results of the few works dealing with this type of classification showed that Multinomial Naïve Bayes (MNB) had the best performance. Al Shboul et al. [26], applied the supervised approach using different classifiers: Decision Table, KNN, SVM, NB and MNB to classify 63,000 reviews based on a 5-star rating system. The results showed that MNB outperforms all others algorithms in terms of accuracy, recall and precision.

However, the Lexicon-based approach needs further improvement to enhance performance on Arabic sentiment classification. The first main factor behind the modest efficiency of this approach is the lack of sentiment dictionaries that can handle the different challenges of the Arabic language. Dictionaries are created using Machine Translation of existing English dictionaries [29] which limits its usefulness for dialects texts, or using seed words [13] which usually contain undesirable noise. Another factor that affects the performance of the Lexicon-based approach is the need of efficient parser to address the sense ambiguities of some words and consider valence shifters such as negation and intensifiers.

## 7. Conclusion

Our objective in this paper was to review the major works that have been devoted to Sentiment Analysis in Arabic. This review clearly demonstrated that Arabic Sentiment Analysis has become one of the research areas that have been drawn the attention of many researchers. Analysis of these works showed that three types of approaches, namely supervised, unsupervised and hybrid are used to tackle a variety of Sentiment Analysis tasks. Nevertheless, Aspect Based Sentiment Analysis, opinion holder extraction and opinion spam detection are the least analysed tasks. Concerning the achieved results, it was revealed that despite the fact that most of the studies reported that their findings are interesting, many challenges need to be sorted out so as to design an effective and mature Sentiment Analysis system. Most of these challenges are inherited from the nature of the Arabic language itself.

## References

- [1] Liu B. Sentiment analysis and opinion mining. *Synth. Lect. Hum. Lang. Technol.* 2012;5(1):1–167.
- [2] K.R. Beesley, Finite-state morphological analysis and generation of Arabic at Xerox Research: Status and plans in 2001, Presented at the ACL Workshop on Arabic Language Processing: Status and Perspective, 2001, vol. 1, pp. 1–8.
- [3] T. Buckwalter, Buckwalter Arabic Morphological Analyzer Version 1.0, 2002.
- [4] Habash NY. Introduction to Arabic natural language processing. *Synth. Lect. Hum. Lang. Technol.* 2010;3(1):1–187.
- [5] N. Habash, R. Eskander, A. Hawwari, A morphological analyzer for Egyptian Arabic, Presented at the Proceedings of the Twelfth Meeting of the Special Interest Group on Computational Morphology and Phonology, 2012, pp. 1–9.
- [6] R.M. Duwairi, R. Marji, N. Sha'ban, S. Rushaidat, Sentiment analysis in Arabic tweets, Presented at the 5th international conference on Information and Communication Systems (icics), 2014, pp. 1–6.
- [7] M. Abdul-Mageed, M.T. Diab, M. Korayem, Subjectivity and sentiment analysis of modern standard Arabic, Presented at the Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers-Volume 2, 2011, pp. 587–591.
- [8] Abdul-Mageed M, Diab M, Kübler S. SAMAR: subjectivity and sentiment analysis for Arabic social media. *Comput. Speech Lang.* 2014;28(1):20–37.
- [9] M. Nabil, M. Aly, A.F. Atiya, ASTD: Arabic sentiment tweets dataset, Presented at the Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, pp. 2515–2519.
- [10] A. Mourad, K. Darwish, Subjectivity and sentiment analysis of modern standard Arabic and Arabic microblogs, Presented at the Proceedings of the 4th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, 2013, pp. 55–64.
- [11] A. Mountassir, H. Benbrahim, I. Berrada, An empirical study to address the problem of unbalanced data sets in sentiment classification, Presented at the

- IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2012, pp. 3298–3303.
- [12] Rushdi-Saleh M, Martín-Valdivia MT, Ureña-López LA, Perea-Ortega JM. OCA: opinion corpus for Arabic. *J. Am. Soc. Inf. Sci. Technol.* 2011;62(10):2045–54.
  - [13] Abbasi A, Chen H, Salem A. Sentiment analysis in multiple languages: feature selection for opinion classification in Web forums. *ACM Trans. Inf. Syst. TOIS* 2008;26(3):12.
  - [14] M. Elhawary, M. Elfeky, Mining Arabic business reviews, Presented at the IEEE International Conference on Data Mining Workshops (ICDMW), 2010 pp. 1108–1113.
  - [15] L. Albraheem, H.S. Al-Khalifa, Exploring the problems of sentiment analysis in informal Arabic, Presented at the Proceedings of the 14th International Conference on Information Integration and Web-based Applications & Services, 2012, pp. 415–418.
  - [16] A. Mountassir, H. Benbrahim, I. Berrada, Sentiment classification on Arabic corpora: preliminary results of a cross-study, 3 E Sémin. Veille Strat. Sci. Technol. VSSST'12, 2012.
  - [17] A. Shoukry, A. Rafea, Sentence-level Arabic sentiment analysis, Presented at the International Conference on Collaboration Technologies and Systems (CTS), 2012, pp. 546–550.
  - [18] S. Oraby, Y. El-Sonbaty, M.A. El-Nasr, Exploring the effects of word roots for Arabic sentiment analysis, Presented at the IJCNLP, 2013, pp. 471–479.
  - [19] M.A. Aly, A.F. Atiya, LABR: a large scale Arabic book reviews dataset, Presented at the ACL (2), 2013, pp. 494–498.
  - [20] S. Ahmed, M. Pasquier, G. Qadah, Key issues in conducting sentiment analysis on Arabic social media text, Presented at the 9th International Conference on Innovations in Information Technology (IIT), 2013, pp. 72–77.
  - [21] M. Al-Kabi, N.M. Al-Qudah, I. Alsmadi, M. Dabour, H. Wahsheh, Arabic/English sentiment analysis: an empirical study, Presented at the Fourth International Conference on Information and Communication Systems (ICICS 2013), 2013.
  - [22] Hamouda AE-DA, El-taher FE. Sentiment analyzer for Arabic comments system. *Int. J. Adv. Comput. Sci. Appl.* 2013;4(3).
  - [23] Hamouda SB, Akaichi J. Social networks' text mining for sentiment classification: the case of Facebook's statuses updates in the "Arabic Spring" era. *Int. J. Appl. Innov. Eng. Manag.* 2013;2(5):470–8.
  - [24] Omar N, Albared M, Al-Shabi A, Al-Moslimi T. Ensemble of classification algorithms for subjectivity and sentiment analysis of Arabic customers' reviews. *Int. J. Adv. Comput. Technol.* 2013;14(5):77–85.
  - [25] S.R. El-Beltagy, A. Ali, Open issues in the sentiment analysis of Arabic social media: a case study, Presented at the 9th International Conference on Innovations in information technology (iit), 2013, pp. 215–220.
  - [26] Al-Kabi MN, Alsmadi IM, Gigieh AH, Wahsheh HA, Haidar MM. Opinion mining and analysis for Arabic language. *IJACSA Int. J. Adv. Comput. Sci. Appl.* 2014;5(5):181–95.
  - [27] Duwairi R, El-Orfali M. A study of the effects of preprocessing strategies on sentiment analysis for Arabic text. *J. Inf. Sci.* 2014;40(4):501–13.
  - [28] R.M. Duwairi, I. Qarqaz, Arabic sentiment analysis using supervised classification, Presented at the International Conference on Future Internet of Things and Cloud (FiCloud), 2014, pp. 579–583.
  - [29] J. Akaichi, Sentiment classification at the time of the Tunisian uprising: machine learning techniques applied to a new corpus for Arabic language, Presented at the Network Intelligence Conference (ENIC), 2014 European, 2014, pp. 38–45.
  - [30] Duwairi R, Ahmed NA, Al-Rifai SY. Detecting sentiment embedded in Arabic social media—a lexicon-based approach. *J. Intell. Fuzzy Syst.* 2015;29(1):107–17.
  - [31] Aldayel HK, Azmi AM. Arabic tweets sentiment analysis—a hybrid scheme. *J. Inf. Sci.* 2015. p. 0165551515610513.
  - [32] R.M. Duwairi, Sentiment analysis for dialectal Arabic, Presented at the 6th International Conference on Information and Communication Systems (ICICS), 2015, pp. 166–170.
  - [33] B. Al Shboul, M. Al-Ayyoub, Y. Jararweh, Multi-way sentiment classification of Arabic reviews, Presented at the 6th International Conference on Information and Communication Systems (ICICS), 2015, pp. 206–211.
  - [34] A. Bayoudhi, L. H. Belguith, H. Ghorbel, Sentiment classification of Arabic documents: experiments with multi-type features and ensemble algorithms, 2015.
  - [35] Elawady RM, Barakat S, El-Bakry HM, Elrashidy N. Sentiment analysis for Arabic and English datasets. *IJICIS* 2015;15(1).
  - [36] E. Refaee, V. Rieser, Benchmarking machine translated sentiment analysis for Arabic tweets, Presented at the NAACL-HLT 2015 Student Research Workshop (SRW), 2015, p. 71.
  - [37] ElSahar H, El-Beltagy SR. Building large Arabic multi-domain resources for sentiment analysis. In: *Computational Linguistics and Intelligent Text Processing*. Springer; 2015. p. 23–34.
  - [38] W. Cherif, A. Madani, M. Kissi, A new modeling approach for Arabic opinion mining recognition, Presented at the Intelligent Systems and Computer Vision (ISCV), 2015, pp. 1–6.
  - [39] H.S. Ibrahim, S.M. Abdou, M. Gheith, Sentiment analysis for modern standard Arabic and colloquial, *ArXiv Prepr. ArXiv:150503105*, 2015.
  - [40] Al-Ayyoub M, Nuseir A, Kanaan G, Al-Shalabi R. Hierarchical classifiers for multi-way sentiment analysis of Arabic reviews. *Int. J. Adv. Comput. Sci. Appl.* 2016;7(2):531–9.
  - [41] A.M. Rabab'ah, M. Al-Ayyoub, Y. Jararweh, M.N. Al-Kabi, Evaluating sentiment strength for Arabic sentiment analysis, Presented at the 7th International Conference on Computer Science and Information Technology (CSIT), 2016, pp. 1–6.
  - [42] S.R. El-Beltagy, T. Khalil, A. Halaby, M. Hammad, Combining lexical features and a supervised learning approach for Arabic sentiment analysis, 2016.
  - [43] H. Al-Rubaiee, R. Qiu, D. Li, Identifying Mubasher software products through sentiment analysis of Arabic tweets, Presented at the 2016 International Conference on Industrial Informatics and Computer Systems (CIICS), 2016, pp. 1–6.
  - [44] Hammad M, Al-awadi M. Sentiment analysis for Arabic reviews in social networks using machine learning. In: *Information Technology: New Generations*. Springer; 2016. p. 131–9.
  - [45] Duwairi RM, Qarqaz I. A framework for Arabic sentiment analysis using supervised classification. *Int. J. Data Min. Model. Manage.* 2016;8(4):369–81.
  - [46] L. Al-Horaibi, M.B. Khan, Sentiment analysis of Arabic tweets using text mining techniques, Presented at the First International Workshop on Pattern Recognition, 2016, pp. 100111F–100111F.
  - [47] L. Abd-Elhamid, D. Elzanfaly, A.S. Eldin, Feature-based sentiment analysis in online Arabic reviews, Presented at the 11th International Conference on Computer Engineering & Systems (ICCES), 2016, pp. 260–265.
  - [48] N. El-Makky et al., Sentiment analysis of colloquial Arabic tweets, 2015.
  - [49] S. Kiritchenko, S.M. Mohammad, M. Salameh, SemEval-2016 Task 7: determining sentiment intensity of English and Arabic phrases, Presented at the Proceedings of the International Workshop on Semantic Evaluation (SemEval), San Diego, California, June 2016.
  - [50] S.R. El-Beltagy, NileTMRG at SemEval-2016 Task 7: deriving prior polarities for Arabic sentiment terms, *Proc. SemEval*, pp. 486–490, 2016.
  - [51] E. Refaee, V. Rieser, iLab-Edinburgh at SemEval-2016 Task 7: a hybrid approach for determining sentiment intensity of Arabic Twitter phrases, *Proc. SemEval*, 2016, pp. 474–480.
  - [52] A. Htaït, S. Fournier, P. Bellot, LSIS at SemEval-2016 Task 7: using web search engines for English and Arabic unsupervised sentiment intensity prediction, *Proc. SemEval*, 2016, pp. 469–473.
  - [53] J.D. Rennie, L. Shih, J. Teevan, D.R. Karger, Tackling the poor assumptions of naive bayes text classifiers, Presented at the ICML, vol. 3, 2003, pp. 616–623.
  - [54] Dodds PS et al. Human language reveals a universal positivity bias. *Proc. Natl. Acad. Sci.* 2015;112(8):2389–94.
  - [55] M. Abdul-Mageed, M.T. Diab, SANA: a large scale multi-genre, multi-dialect lexicon for Arabic subjectivity and sentiment analysis, Presented at the LREC, 2014, pp. 1162–1169.
  - [56] Kouloumpis E, Wilson T, Moore JD. Twitter sentiment analysis: the good the bad and the omg! *ICWSM* 2011;11:538–41.
  - [57] E. Refaee, V. Rieser, An Arabic twitter corpus for subjectivity and sentiment analysis, Presented at the LREC, 2014, pp. 2268–2273.
  - [58] Alhazmi M, Salim N. Arabic opinion target extraction from tweets. *ARPN J. Eng. Appl. Sci.* 2015;10(3).
  - [59] M. Al-Smadi, O. Qawasmeh, B. Talafha, M. Quwaider, Human annotated Arabic dataset of book reviews for aspect based sentiment analysis, Presented at the 3rd International Conference on Future Internet of Things and Cloud (FiCloud), 2015, pp. 726–730.
  - [60] M. Pontiki, D. Galanis, J. Pavlopoulos, H. Papageorgiou, I. Androustopoulos, S. Manandhar, Semeval-2014 task 4: aspect based sentiment analysis, Presented at the Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), 2014, pp. 27–35.
  - [61] I. Obaidat, R. Mohawesh, M. Al-Ayyoub, M. Al-Smadi, Y. Jararweh, Enhancing the determination of aspect categories and their polarities in Arabic reviews using lexicon-based approaches, Presented at the IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT), 2015, pp. 1–6.
  - [62] M. Abdul-Mageed, M.T. Diab, AWATIF: a multi-genre corpus for modern standard Arabic subjectivity and sentiment analysis, Presented at the LREC, 2012, pp. 3907–3914.
  - [63] N.A. Abdulla, N.A. Ahmed, M.A. Shehab, M. Al-Ayyoub, Arabic sentiment analysis: Lexicon-based and corpus-based, Presented at the IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT), 2013 pp. 1–6.
  - [64] M. Al-Kabi, A. Gigieh, I. Alsmadi, H. Wahsheh, M. Haidar, An opinion analysis tool for colloquial and standard Arabic, Presented at the Proceedings of the 4th International Conference on Information and Communication Systems, ICICS, 2013, vol. 13.
  - [65] N. Abdulla, R. Majdalawi, S. Mohammed, M. Al-Ayyoub, M. Al-Kabi, Automatic lexicon construction for Arabic sentiment analysis, Presented at the International Conference on Future Internet of Things and Cloud (FiCloud), 2014, pp. 547–552.
  - [66] Badaro G, Baly R, Hajj H, Habash N, El-Hajj W. A large scale Arabic sentiment lexicon for Arabic opinion mining. *ANLP* 2014;165:2014.
  - [67] A. Bayoudhi, H. Koubaa, H. Ghorbel, L. Hadrich Belguith, Vers un lexique arabe pour l'analyse des opinions et des sentiments, Presented at the Proc of the 5th International Conference on Arabic Language Processing CITALA, vol. 14, 2014.
  - [68] Mahyoub FH, Siddiqui MA, Dahab MY. Building an Arabic sentiment lexicon using semi-supervised learning. *J. King Saud Univ. – Comput. Inf. Sci.* 2014;26(4):417–24.
  - [69] H.S. Ibrahim, S.M. Abdou, M. Gheith, Idioms-proverbs lexicon for modern standard Arabic and colloquial sentiment analysis, *ArXiv Prepr. ArXiv:150601906*, 2015.

- [70] A. Assiri, A. Emam, H. Al-Dossari, Saudi twitter corpus for sentiment analysis, *World Acad. Sci. Eng. Technol. Int. J. Comput. Electr. Autom. Control Inf. Eng.*, vol. 10, no. 2, pp. 180–183.
- [71] R. Eskander, O. Rambow, SLSA: a sentiment Lexicon for standard Arabic.
- [72] [Al-Ayyoub M, Essa SB, Alsmadi I. Lexicon-based sentiment analysis of Arabic tweets. Int. J. Soc. Netw. Min. 2015;2\(2\):101–14.](#)
- [73] M. Elarnaoty, S. AbdelRahman, A. Fahmy, A machine learning approach for opinion holder extraction in Arabic language, *ArXiv Prepr. ArXiv:12061011*, 2012.
- [74] [Dahab M, Al Asmari S. Towards discovering the sentiment holder in Arabic text. Int. J. Comput. Appl. 2015;111\(16\).](#)
- [75] H.A. Wahsheh, M.N. Al-Kabi, I.M. Alsmadi, SPAR: a system to detect spam in Arabic opinions, Presented at the IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT), 2013, pp. 1–6.
- [76] [Wahsheh HA, Al-Kabi MN, Alsmadi IM. A link and content hybrid approach for Arabic web spam detection. Int. J. Intell. Syst. Appl. 2012;5\(1\):30.](#)
- [77] [Abu Hammad A, El-Halees A. An approach for detecting spam in Arabic opinion reviews. Int. Arab J. Inf. Technol. IAJIT 2015;12\(1\).](#)
- [78] [Medhat W, Hassan A, Korashy H. Sentiment analysis algorithms and applications: a survey. Ain Shams Eng. J. 2014;5\(4\):1093–113.](#)

**Naaima Boudad** is a PhD student at ENSIAS, Mohammed V University, Rabat, Morocco. She received her Engineering degree at INPT in 2005. Her fields of interest are Text mining, Data mining, and Artificial Intelligence.

**Rdouan Faizi** is an Associate Professor at ENSIAS, Mohammed V University, Rabat, Morocco.

**Rachid Oulad Haj Thami** is a Full Professor at ENSIAS, Mohammed V University, Rabat, Morocco.

**Raddouane Chiheb** is an Associate Professor at ENSIAS, Mohammed V University, Rabat, Morocco.