



Article

# A New Ontology-Based Method for Arabic Sentiment Analysis

Safaa M. Khabour <sup>1</sup>, Qasem A. Al-Radaideh <sup>1</sup> and Dheya Mustafa <sup>2,\*</sup>

<sup>1</sup> Department of Information Systems, Faculty of Information Technology and Computer Sciences, Yarmouk University, Irbid 21163, Jordan; safaa.khabour@ses.yu.edu.jo (S.M.K.); qasemr@yu.edu.jo (Q.A.A.-R.)

<sup>2</sup> Department of Computer Engineering, Faculty of Engineering, The Hashemite University, Zarqa 13133, Jordan

\* Correspondence: dheya@hu.edu.jo

**Abstract:** Arabic sentiment analysis is a process that aims to extract the subjective opinions of different users about different subjects since these opinions and sentiments are used to recognize their perspectives and judgments in a particular domain. Few research studies addressed semantic-oriented approaches for Arabic sentiment analysis based on domain ontologies and features' importance. In this paper, we built a semantic orientation approach for calculating overall polarity from the Arabic subjective texts based on built domain ontology and the available sentiment lexicon. We used the ontology concepts to extract and weight the semantic domain features by considering their levels in the ontology tree and their frequencies in the dataset to compute the overall polarity of a given textual review based on the importance of each domain feature. For evaluation, an Arabic dataset from the hotels' domain was selected to build the domain ontology and to test the proposed approach. The overall accuracy and f-measure reach 79.20% and 78.75%, respectively. Results showed that the approach outperformed the other semantic orientation approaches, and it is an appealing approach to be used for Arabic sentiment analysis.

**Keywords:** sentiment analysis; arabic language; domain ontology; semantic orientation; feature level sentiment analysis; features selection and weighting; domain features



**Citation:** Khabour, S.M.; Al-Radaideh, Q.A.; Mustafa, D. A New Ontology-Based Method for Arabic Sentiment Analysis. *Big Data Cogn. Comput.* **2022**, *6*, 48. <https://doi.org/10.3390/bdcc6020048>

Academic Editors: S. Ejaz Ahmed, Shuangge Steven Ma and Peter X.K. Song

Received: 1 April 2022  
Accepted: 22 April 2022  
Published: 29 April 2022

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



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The Web offers a massive virtual space where users can express and publish their opinions and experiences. People use social media daily as a primary place in a wide range of applications in their lives, not only for social life purposes but also e-learning, e-commerce, politics, and many other applications. In the Middle East (where Arabic is the mother language), Facebook and Twitter were determined as the most prevalent social media websites that affect youth [1,2]. While web content has witnessed an unprecedented increase in size, the process of extracting useful information is becoming more challenging as well [3,4].

Sentiment analysis, or opinion mining, is a type of text mining research that depends mainly on Machine Learning (ML) and Natural Language Processing (NLP) approaches for mining subjective texts [4–9]. Sentiment analysis research scope in the field of computer science is rising very quickly [10]. The semantic web is a logical expansion of the World Wide Web, which is intended to make the web more machine-understandable [11]. The ontology is an essential semantic technology used widely for data handling in the semantic web [12,13]. Ontologies facilitate communication between humans and agents; they also describe the domain theories for the explicit representation of the semantics of the data [14] and web interoperability [15]. Ontology is a systematic account of existence [16], where it can be used to formalize and model-specific domain knowledge to be represented and applied in different fields, such as the semantic web, artificial intelligence, system engineering, information architecture, enterprise bookmarking and biomedical informatics. Furthermore, the ontology concept is valuable in text mining applications, such as in-text

clustering and classification [9,17], and in-text summarization [18], which can be applied to different domains.

Ontology is usually used for accomplishing two main tasks in sentiment analysis research, either for lexicon creation or for aspect (feature) extraction [19–26]. Sentiment analysis researchers are immediately directed toward ontology-based approaches to represent a common sense knowledge base [27,28]. Tartir and Abdul-Nabi [29] utilized ontology for lexicon creation as they created Arabic Sentiment Ontology (ASO). SenticNet is another example of a concept-based resource, which was created to comprise 5732 single and multi-word concepts along with their polarity scores in the range of  $-1$  to  $+1$  [30]. The use of ontology in such text mining applications has achieved considerable results [31].

Recently, few studies were conducted on sentiment analysis for the Arabic language as compared with those for the English language; several researchers shifted towards the analysis of the Arabic language [32]. Major challenges of Arabic sentiment analysis (ASA) are related to the language nature [33]. The rich morphology of the Arabic language, with which you can express the same meaning in different ways by combining different stems, roots, prefixes, and suffixes. This increases the need to conduct a morphological analysis, where each term is divided into morphemes, and each morpheme combines with morphological information such as root, stem, POS, and affix [5]. This complex challenge raises the need to improve convenient NLP tools to handle morphological analysis, tokenization, stemming, spell checking, lemmatization, part-of-speech tagging, and pattern matching [34,35]. Furthermore, the existence of the various speaking dialects in real life and on the web [8], the rich Arabic synonyms, and the low number of studies that are concentrated on ASA using domain-specific ontologies and features importance, are considered motivations for this research.

This study used ontology to extract features based on ontology concepts, along with concepts' levels and concepts' frequencies. Then, it determines the importance of each feature based on its location in the ontology tree and its appearance in the review's dataset. The proposed semantic orientation approach benefits from domain ontologies and sentiment lexicons for accomplishing ASA to increase the accuracy of the analysis process of the users' subjective opinions. In this approach, we used ontology to explain some knowledge about the domain features during the feature extraction and selection phases.

We intend to solve the limitations in the previous research on ASA focused on the semantic orientation approaches, where the semantic features in the proposed approach are treated with their different weights of importance in the subjective text. Overall, this paper makes the following contributions:

1. Building a semantic orientation approach using ontology for mining the different opinions to decrease the effort needed by ordinary users or organizations to make more accurate sentiments classification. The approach is working at the level of semantic features, which are extracted and weighted using the domain ontology.
2. Using the domain features' levels to determine the polarity of the overall review. Also, the important domain features from the users' point of view are used to efficiently calculate the overall semantic polarity of a subjective text. This approach is different from the previous ontology-based approaches in using a weighting method with two factors to identify the different weights of importance for each semantic domain feature.
3. Evaluating the proposed approach with an Arabic dataset from the hotels' domain, which was selected to build the domain ontology.

The rest of the paper is organized as follows: The next section discusses related work. Section 3 presents the research methodology and proposed sentiment analysis approach. Evaluation and results are presented and discussed in Section 4, followed by conclusions in Section 5.

## 2. Related Work

Sentiment classification approaches can be categorized into three fields: Sentimental SO, Machine Learning (ML), and hybrid approaches. Nithish et al. [36] proposed a feature-based sentiment analysis model of the English language using product reviews. They applied the feature level analysis to mobile product reviews and reached 70% accuracy. Thakor and Sasi [19] presented a sentiment analysis approach to classifying negative sentiments in social media content based on ontology. The proposed approach successfully classified 253 negative tweets out of 494 tweets. In [21], the authors proposed a sentiment analysis approach based on Latent Dirichlet Allocation (LDA) topic clusters, domain ontology, and SentiWordNet for Nokia 6610 cellular phone reviews. The precision of the extracted product features was 76.1%.

Alfonso and Sardinha [22] proposed an approach for holding the relationships between aspects, associations of aspects, and their expressions of opinion for aspect-based sentiment analysis using a fuzzy ontology. They tested their approach on the hotels' domain, where each aspect of the hotel got a score, and then they calculated the total score for the hotel by accumulating the scores of each aspect.

Zehra et al. [23] proposed an approach to construct a recommendation system based on sentiment analysis using ontology. The researchers focused on a Facebook closed group that includes posts and comments about various schools collected randomly. Salas-Zárate et al. [24] proposed an aspect-level opinion mining approach to the diabetes domain using ontologies to identify the aspects related to diabetes in the tweets.

Lazhar and Yamina [37] examined the effectiveness of domain ontologies in ASA. Mahyoub et al. [38] presented in their study a sentiment lexicon for the Arabic language when the proposed system worked to specify the sentiment scores for each word included in the Arabic WordNet. The accuracy of the classification reached 96%. Soliman et al. [39] presented an approach to building a Slang Sentimental Words and Idioms Lexicon (SSWIL) of opinion words. They also worked to categorize Arabic news comments on Facebook separating the SVM classifier into two classes: satisfy and dissatisfy, with an accuracy rate of 86.86%.

ML approaches use ML techniques such as Naive Bayes (NB), Support Vector Machines (SVM), Bayesian Network (BN), Maximum Entropy (ME), and Neural Network (NN) for building classifier models [12,40–46]. Several ML approaches were proposed for sentiment analysis of standard or dialect Arabic tweets dataset based on classes of polarity [47–52] or using the rough set-based concepts [53]. Tripathy et al. [54] presented an ML approach for English language sentiment classification. Jagdale et al. [55] applied sentiment analysis to the English language using ML techniques on a dataset collected from Amazon about different products reviews. Other studies considered sentiment analysis for different languages, such as Chinese [56], Turkish [57], and Lithuanian [58].

Hybrid approaches combine different semantic orientation approaches with different machine learning approaches to improve the results of the sentiment analysis process [59–61]. Several studies proposed a hybrid approach for sentiment analysis of different Arabic dialects tweets [62–64] and tweets of product reviews [65].

We benefit from these studies to build an enhanced approach for ASA using the ontology model. Tartir and Abdul-Nabi [29] focused on the semantic relations between sentiments and their instances to present a semantic orientation approach. In other semantic orientation approaches such as the studies of Thakor and Sasi [19–24], they focused on the use of ontology for feature identification and extraction without considering any other information from the ontology tree such as the levels of features, while El-Halees and Al-Asmar [25] used the levels of features to calculate the polarity, by multiplying each feature level with its sentiment polarity, where the levels indicate the feature importance. In this research, we used the ontology to identify and extract the domain features and their levels, while at the same time the frequencies of these features in the review's dataset are also used to identify the importance of each feature.

### 3. Method

This section presents and discusses the methodology followed in this study. The first subsection describes the overall approach design. The Arabic resources used in this work are described in Section 3.2, while the third subsection describes the main research phases and the entire steps in each phase in more detail.

#### 3.1. Overall Approach Design

The overall methodology to classify Arabic textual reviews based on sentiment analysis using ontology is divided into five main phases, and each phase has several steps, as illustrated in Figure 1.

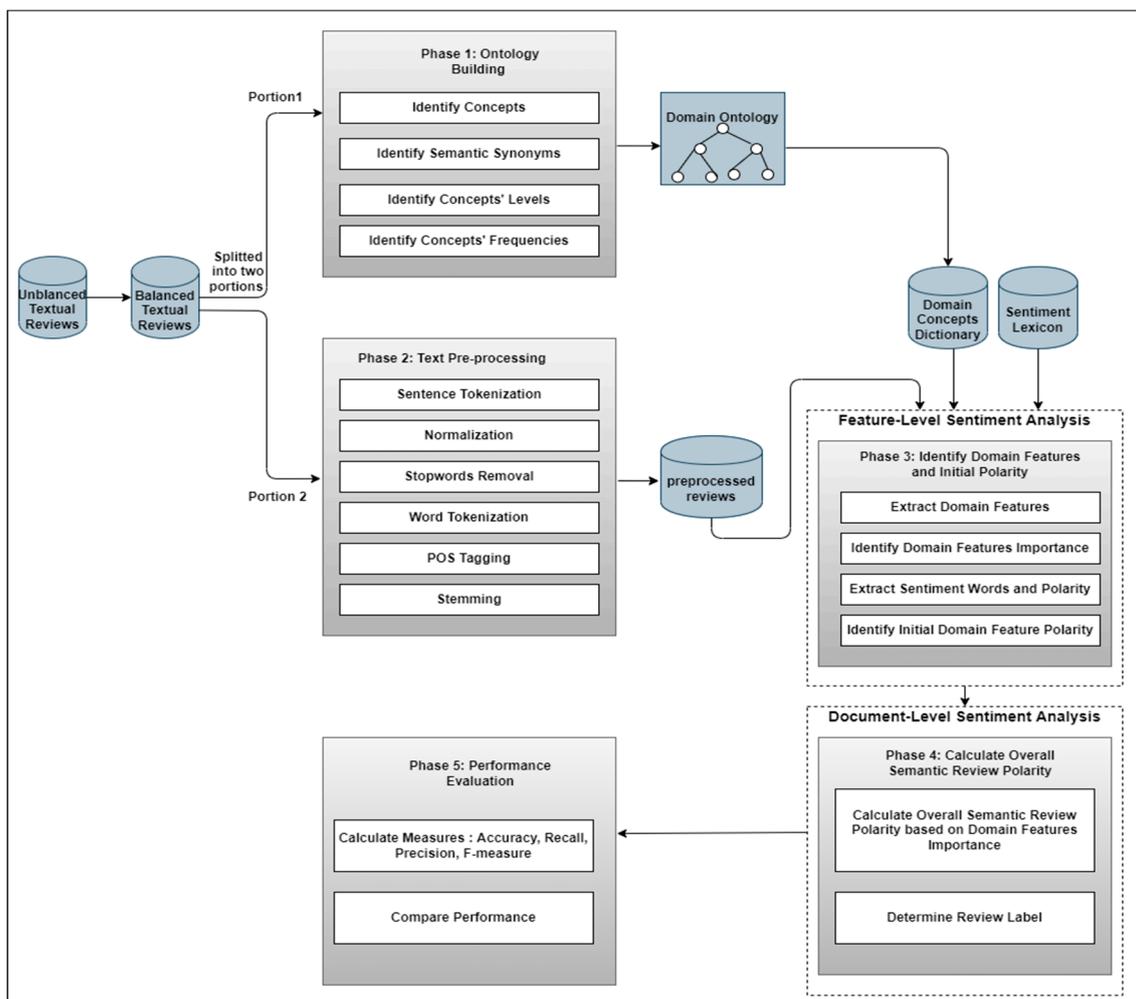


Figure 1. Overall approach design.

#### 3.2. Description of the Arabic Resources

To illustrate the steps of the proposed method, it is beneficial to introduce the Arabic resources that were used in the evaluation; this will help the reader to gain insight into the proposed method. We used ElSahar and El-Beltagy [66] dataset to extract the domain-specific ontology and to evaluate and test the model. The overall dataset comprises around 33 thousand automatically annotated reviews in various domains which are: movies, restaurants, hotels, and products. Also, domain-specific lexicons contain about two thousand entries semi-automatically generated from the reviews.

The hotel reviews dataset contains around 15 thousand Arabic user reviews, extracted from the TripAdvisor website. The authors employed the open-source Scrapy framework,

for establishing custom web crawlers. Table 1 describes the general statistics of the hotels' datasets of ElSahar and El-Beltagy [66]. Table 2 holds a sample hotel review from the dataset, where each row is considered as a user opinion on a particular hotel and the identified polarity for that review. We added the review translation.

**Table 1.** Hotel reviews' statistics of ElSahar and El-Beltagy dataset.

Total Positive Reviews	Total Negative Reviews	Total Neutral Reviews
10,775	2647	2150

**Table 2.** Sample hotel review from ElSahar and El-Beltagy dataset.

#	Review Text	Polarity
1	<p>ممتاز- رائع-انصح الجميع به في شهر اكتوبر 2013 نزلت 3 يوم في الفندق-كان اكثر من رائع-اسعار مناسبة-خدمة ممتازة-الغرف نظيفة . جدا والحمامات رائعة-ادارة الفندق والموظفين اكثر من ممتازة-حراس الامن ممتازة-المطعم والبار والمطهي رائعة حقا-لم ارى افضل من ذلك الفندق في اديس ابابا.</p> <p><b>Translation:</b> Excellent-wonderful-I recommend it to everyone. In October 2013 I stayed 3 days in the hotel-it was more than wonderful-suitable prices-excellent service-rooms are very clean and bathrooms are wonderful-hotel management and staff are more than excellent-security guards are excellent-the restaurant, bar and nightclub are wonderful Really-I've never seen a better hotel than that in Addis Ababa.</p>	1 (Positive)

Previous ASA studies suffered from the unavailability of adequate resources that classify the opinion words (sentiment lexicons). Although there exist some efforts to build lexicons in Arabic, they still have limitations such as unclear usability, small size, and non-publicly shared lexicons. ArSenL is an Arabic SentiWordNet lexicon developed by Badaro et al. [67] to solve the previously mentioned limitations. The developers created the first large-scale publicly shared resource for opinion mining in standard Arabic. Their lexicon was built based on three different available resources: English sentiWordNet, the Standard Arabic Morphological Analyzer (SAMA), and Arabic WordNet.

Two values are attached with each existing lemma entry in the lexicon which indicates the positive and negative polarity scores. It contains four types of Part of Speech (POS) tags (adjective, noun, verb, and adverb). The lemmas are presented in Buckwalter's (2004) format to facilitate the NLP processes. ArSenL contains a total of around 28,760 lemmas and 157,969 Synsets which is considered a large-scale Arabic sentiment lexicon. Table 3 provides a sample of the ArSenL lexicon content; we added a column that represents each sentiment in an Arabic form and its translation in English as well.

**Table 3.** Sample of ArSenL lexicon.

Arabic Form	Meaning	Aramorph Lemma	POS	Positive Sentiment Score	Negative Sentiment Score
فرح	Happyness	faraH_1	Noun	0.5	0.125
حزن	Felt sad	Hazin-a_1	Verb	0	0.5
شريف	Honorable	\$ariyf_2	Adjective	1	0

### 3.3. Main Phases of the Approach

This section aims to briefly describe and discuss the main phases depicted in Figure 1 by explaining the steps and processes which are used for each phase.

#### 3.3.1. Ontology Building

For the proposed semantic orientation approach of sentiment analysis, we need to build domain ontology. This ontology is used as a domain concept dictionary to extract the domain features with their importance. In this phase, we built domain ontology

by extracting the concepts that are relevant to the hotel domain using Latent Dirichlet Allocation (LDA) with manual approaches. Two lists of domain concepts are generated; one of them is extracted using the LDA algorithm, and the other list is extracted from the dataset manually because LDA ignores the concepts with low frequencies [26].

Figure 2 provides a graphical representation of LDA topic modeling. The Latent Dirichlet Allocation (LDA) model, proposed by Blei et al. [68], is an unsupervised method that is well-known in text mining applications. It can recognize the latent topics from several documents automatically [26]. LDA is used to arrange a document text into specified topics. It generates topics per documents model and words per topic model, using Dirichlet distributions [69]. Each topic is a collection of keywords, and each keyword participates in a specific weightage to the topic [68]. Variables and parameters which appear in Figure 2 of the LDA model are interpreted as:  $D$  is the number of documents in the corpus,  $N$  is the number of words in a specified document,  $A$  is the Dirichlet prior parameter on the topic distributions per document,  $B$  is the Dirichlet prior parameter on the word distribution per topic,  $\Theta$  is the topic distribution for a specified document,  $\Phi$  is the word distribution for a specified topic  $k$ ,  $TP$  is the topic assignment for a word in the specified document, and  $W$  is the specified word.

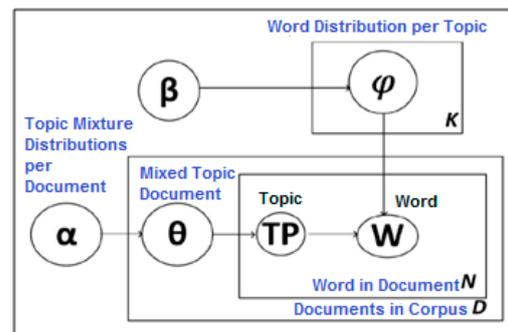


Figure 2. Graphical model representation of LDA.

In the proposed approach, at first, the LDA is used to generate topic clusters from the dataset where each topic contains a group of keywords. To implement the LDA model using Python, Algorithm 1 is used. The portion of the dataset which is assigned for building ontology is imported in Python. Several preprocessing steps are utilized to normalize reviews' sentences, tokenize them into words, and remove unnecessary words. Two inputs are required for running the LDA modeling which are the dictionary and the corpus that report the distinct words and their repetitions in the training data. The Term Frequency-Inverse Document Frequency (TF-IDF) transformation is applied to the entire corpus, and then the LDA is run. The resulting topics contain keywords unlike to be domain concepts such as sentiments [21], so human evaluation is used to filter these topics and to judge each keyword to determine suitable domain concepts. Table 4 provides a sample of LDA-generated topics from the dataset of ElSahar and El-Beltagy [66], where the keywords in bold represent possible domain concepts.

For the manual list of concepts, human evaluators are contributed to extracting domain concepts from a set of reviews manually, and then the extracted concepts are compared with the list of concepts using LDA and combined the two lists. The evaluators read the final list to identify the distinct concepts and their synonyms; also, they identify the relationships between them to determine their positions from the top to the bottom of the ontology tree. The final ontology is presented using the Protégé tool [70] to facilitate identifying the level for each concept, where the classes and subclasses represent concepts and sub-concepts for that domain [71]. We used the Protégé tool only to draw the ontology instead of manual drawing.

**Algorithm 1** Building the LDA topic model**Input:** Hotel Reviews Dataset**Output:** Topics with Keywords

- 1- Load the hotel reviews dataset.
- 2- Preprocess the reviews dataset:
  - Normalization.
  - Tokenization.
  - Stopword Removal.
- 3- Apply Bag-of-Words (BoWs) on the dataset.
- 4- Apply TF-IDF transformation to the entire corpus.
- 5- Train the LDA topic model using the Gensim module.
- 6- Present the topics of the LDA model.

**Table 4.** Sample of the generated LDA topics and keywords.

Topics	Keywords
Topic 1	افضل, قذر, استقبال, افطار, لطيف, توجد, يمكن, طاقم, موظف, رائع, جيد, مطعم <b>Restaurant</b> , good, wonderful, <b>employee</b> , <b>staff</b> , can, there, nice, <b>breakfast</b> , <b>reception</b> , dirty, better, ... . . .
Topic 2	اسوء, يتوفر, موقع, نظيف, كبير, حمام, سرير, مريح, شاطئ, جميل, طعام, ممتاز Excellent, <b>food</b> , beautiful, <b>beach</b> , comfortable, <b>bed</b> , <b>bathroom</b> , large, clean, <b>location</b> , available, worst, ... . . .

After identifying the concepts, the Arabic WordNet browser and Google translation are used to search for more semantic Arabic synonyms for each concept. This phase aims to extract all semantic domain features and all words that have the same meaning as the domain features. Table 5 provides an example of semantic synonyms for extracted hotel concepts from the dataset of ElSahar and El-Beltagy [66]. Table 6 shows the total number of distinct domain concepts and the total number of levels in the constructed hotel ontology.

**Table 5.** Example of semantic synonyms for extracted hotel concepts.

No	Concept	Meaning	Semantic Synonyms						
1	فندق	Hotel	نزل	خان	اوتيل	هوتيل	بنسيون		
2	طاقم	Staff	كادر	فريق	ستاف	تيم			
3	غطاء	Blanket	لحاف	بطانية	دثار	ملاءة	ملاية	حرام	شرشف

**Table 6.** Characteristics of the constructed hotel ontology model.

Number of Distinct Concepts	Number of Ontology Levels
203	6

For each concept, the level is identified using the Protégé structure, we assume that the highest level (6) is at the ontology tree root and the lowest level (1) is at the last bottom feature in the ontology tree. Furthermore, for each concept, we identify the total frequency by calculating the sum of the concept's frequency and its synonyms' frequencies, and then two important weights are calculated for each of them. All the needed information from the ontology is stored in a separate file as a domain concepts dictionary. Each row in the domain concepts dictionary consists of Domain\_Concept, Concept\_Level\_Importance, Concept\_Frequency\_Importance, and List\_of\_Synonyms.

### 3.3.2. Text Preprocessing

The reviews dataset is unstructured and contains stopwords, so it needs to be preprocessed. Text preprocessing is intended to make the reviews consistent and to represent them in some standard form to facilitate conducting systematic processes. Some NLP processes were used to preprocess the textual reviews. These processes include sentence tokenization, normalization, stopword removal, word tokenization, POS tagging, and stemming. Table 7 provides an example for each of them. English translation of the Arabic input is added.

**Table 7.** Example of a user review after each of the preprocessing steps.

Original Review		Input	
		موقع مركزي في المدينة! كذلك السعر مناسب جدا لجميع الفئات و الغرف نظيفة.	
		Translation: Central Location in the City! Also, the Price Is Very Suitable for All Categories and the Rooms Are Clean.	
No	Preprocessing Step	Output	
1	Sentence Tokenization	Central location in the city! Also, the price is very suitable for all categories and the rooms are clean.	- موقع مركزي في المدينة! - كذلك السعر مناسب جدا لجميع الفئات و الغرف نظيفة.
2	Normalization	Central location in the city Also, the price is very suitable for all categories and the rooms are clean	- موقع مركزي في المدينة - كذلك السعر مناسب جدا لجميع الفئات و الغرف نظيفة
3	Stopwords Removal	Central location the city Also, the price is very suitable for all categories the rooms are clean	- موقع مركزي المدينة - السعر مناسب جدا لجميع الفئات الغرف نظيفة
4	Word Tokenization	"location", "central", "the city" "the price", "very", "suitable", "for all" "categories", "the rooms", "clean"	- "موقع", "مركزي", "المدينة" - "السعر", "مناسب", "جدا", "لجميع", "الفئات", "الغرف", "نظيفة"
5	POS Tagging (using the Stanford Tagger)	[('موقع', 'NN'), ('مركزي', 'JJ'), ('المدينة', 'DTNN')] [('السعر', 'DTNN'), ('مناسب', 'JJ'), ('جدا', 'NN'), ('جميع', 'NN'), ('الفئات', 'DTNNS'), ('الغرف', 'DTNN'), ('نظيفة', 'JJ')]	
6	Light Stemming (using Tashaphyne)	"location", "center", "city" "price", "very", "suitable", "all" "category", "rooms", "clean"	- ["موقع", "مركز", "مدین"] - ["سعر", "مناسب", "جد", "جميع", "فى", "غرف", "نظيف"]

### 3.3.3. Domain Features and Initial Polarity Identification

This phase aims to distinguish the domain features and sentiment words using the POS, where the nouns are considered as candidate domain features for identification and extraction using the domain dictionary. The noun tags using the Stanford POS tagger [72] are NN, DTNN, NNP, DTNNP, NNS, DTNNS, NNPS, DTNNPS, NOUN, NOUN\_QUANT. The other words such as adjectives, verbs, and the residual nouns which were not found in the domain dictionary are considered candidate sentiment words to match with the lexicon [37].

To extract the sentiment words around each domain feature, the N-gram-around method achieves considerable results in identifying the sentiment words related to each domain feature [20,24]. The initial polarity for the domain feature is calculated based on the sum of the positive scores and the sum of the negative scores for the sentiment words which are extracted using the N-gram-around method.

To search and match each sentiment word with the lexicon, three methods are used: the original word is matched with the lexicon; if not found, the word stem is matched with the lexicon; and if not found, the word root is matched with the lexicon. If neither the word nor its stem nor its root is found, its sentiment polarity is considered zero. For this step, we used the Tashaphyne stemmer [73], which is supported in Python, to generate both stems and roots.

Negations and intensifiers are handled during identifying sentiment words' polarities. Negation in the Arabic language is expressed by adding (عدم/لن/لا/مش) "not" before a verb, noun, or adjective. If any of the negation terms appears before a sentiment word, it counters the meaning of that word; adding a negation particle before a positive word would make it negative, and vice versa. For example, in the sentence (الادارة مش أمانة), the word (أمانة) is positive and its positive and negative scores using the ArSenL lexicon are (0.083, 0.05), respectively. When the negation particle (مش) comes before it, its scores change to (0.05, 0.083), which is negative. Intensifiers in the Arabic language, such as (تماما/كثيرا جدا) are added after a sentiment word to emphasize the meaning and indicate the strength of the meaning. So, we consider that when they appear after a sentimental word, the polarity for that word is doubled. For example, (الفندق رائع) is a sentiment word with positive and negative scores of (0.402, 0.069), respectively. After adding (جدا) to the sentence, its scores changed to (0.804, 0.138). Table 8 provides an example of this phase.

**Table 8.** Example of domain features and initial polarity identification.

Original Review	Input
	موقع رائع في مركز المدينة الا انني شعرت بالاستياء لإهمالهم بسبب وجود سحابة قذرة
Step	Output
Extract Domain Features with Importance	(موقع, NN):(levelImportance= 5, freqImportance= 1) (سحابة, NN):(levelImportance= 2, freqImportance= 0.25)
Extract Around Sentiments	['شعرت', 'المدينة', 'مركز', 'رائع'] (موقع) ['وجود', 'لإهمالهم', 'بالاستياء', 'شعرت', 'قذرة'] (سحابة)
Initial Domain Feature Polarity Identification	Positive Score (موقع) = 0.4282/Negative Score (موقع) = 0.09063 - Initial Polarity (موقع) = 0.4282 – 0.09063 = 0.33757 Positive Score (سحابة) = 0.20594/Negative Score (سحابة) = 0.7226 - Initial Polarity (سحابة) = 0.20594 – 0.7226 = –0.51666

### 3.3.4. Overall Semantic Review Polarity Calculation

Based on the extracted semantic domain features for each review, we need to calculate a total semantic review polarity. The initial polarity of each domain feature is affected by the importance of that feature. The Formula (1) is used to calculate overall semantic review polarity based on semantic features' importance:

$$\text{Overall Semantic Review Polarity} = \sum_{i=1}^n \text{Initial Polarity (DF}_i\text{)} \times (\text{L (DF}_i\text{)} + \text{F (DF}_i\text{)}) \quad (1)$$

where n is the number of the extracted domain features from a review, DF<sub>i</sub> represents the specific domain feature that has an initial polarity, L represents the level of importance of the domain feature (DF<sub>i</sub>), which is identified based on its level in the ontology tree, and F represents the frequency importance of the domain feature (DF<sub>i</sub>) which takes the following values—0.1, 0.25, 0.50, 0.75, 1—to indicate its importance from domain users' point of view. Since features' levels are not dependent on the dataset, we consider the domain feature frequency to represent the importance of the domain features as they are repeated in the dataset. High frequent domain features in the dataset means that users are more interested in those features in that domain than the other ones. Domain features are divided into five groups based on their frequencies; the most frequent features in the dataset get the highest importance value as (1), and so on, whereas the lowest frequent features get the lowest importance value as (0.1). We experiment with assigning different weights for this

factor for each group of domain features. We noticed that these weights of importance have improved the performance of the semantic orientation sentiment analysis.

The review label is determined as positive (+1) if the overall semantic review polarity is greater than or equal to zero because the third class (neutral) is ignored in the proposed approach and we noticed that the number of reviews where their total semantic polarity exactly equals to zero are very few in the dataset, so we considered them as positive reviews. Conversely, the review label is determined as negative (−1) if the overall semantic review polarity is less than zero. Table 9 illustrates the phase of calculating the overall semantic review polarity using the previous phase example.

**Table 9.** Example of calculating overall semantic review polarity.

	Input
Extracted Domain Features and Identified Importance	-Initial Polarity (موقع) = 0.33757 (levelImportance = 5, freqImportance = 1) -Initial Polarity (سجادة) = −0.51666 (levelImportance = 2, freqImportance = 0.25)
Step	Output
Calculating Overall Semantic Polarity for the Review based on Domain Features Importance	(موقع) = Initial Polarity * (L + F) = 0.33757 *(5 + 1) = 2.02542 (سجادة) = Initial Polarity * (L + F) = −0.51666 *(2 + 0.25) = −1.162485 Overall Semantic Review Polarity = 2.02542 + (−1.162485) = 0.862935
Determine Review Label	Positive (+1)

The overall review polarity is considered positive, although the review contains one feature that is considered positive with an initial polarity of (+0.33757), and one feature that is considered negative with an initial polarity of (−0.51666), where the negative feature has the higher initial value. Since the positive feature has higher importance than the negative feature; the total importance of the positive feature is (6) and for the negative feature is (2.25).

### 3.3.5. Performance Evaluation

In this phase, some performance evaluation metrics are used to measure the performance of the proposed approach, and to compare it with some other semantic orientation approaches used by researchers in the literature. The performance evaluation measures are accuracy, recall, precision, and f1-measure. Referring to [54], the precision and recall measures can be computed for the positive class using the following equations:

$$\text{Precision (Positive)} = \frac{TP}{TP + FP} \quad (2)$$

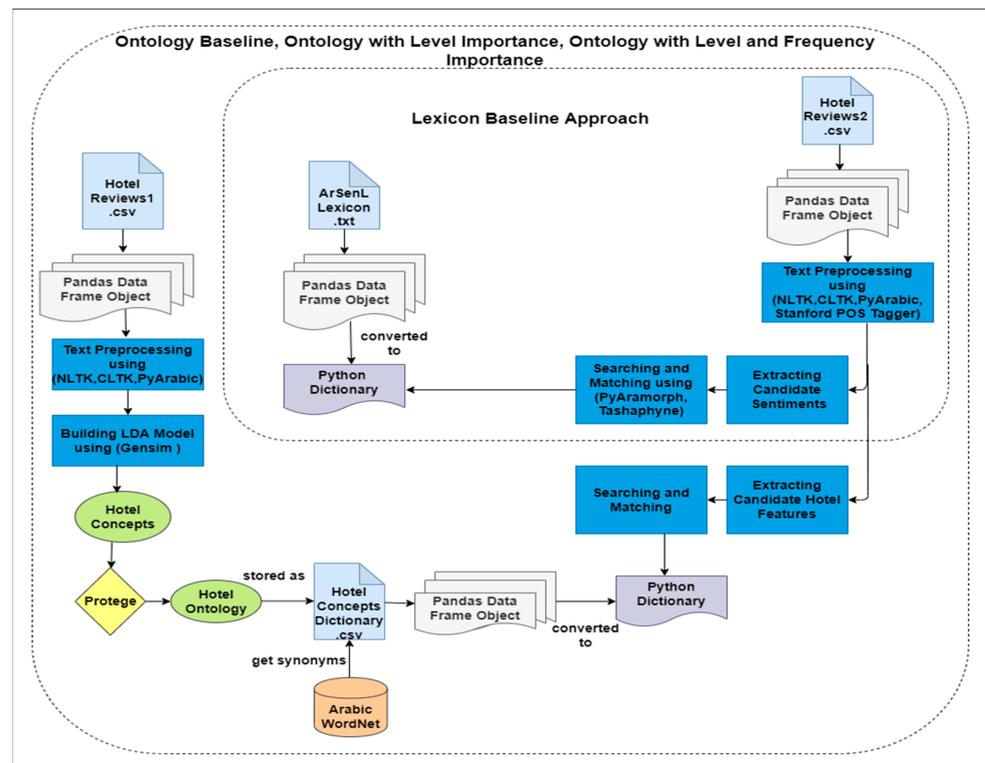
$$\text{Recall (Positive)} = \frac{TP}{TP + FN} \quad (3)$$

where:

- TP (True Positive): represents the number of reviews that are classified as positive in both original classifications and predicted classifications.
- TN (True Negative): represents the number of reviews that are classified as negative in both original classifications and predicted classifications.
- FP (False Positive): represents the number of reviews that are classified as positive in the predicted classifications, while classified as negative in the original classifications.
- FN (False Negative): represents the number of reviews that are classified as negative in the predicted classifications, while classified as positive in the original classifications.

#### 4. Results and Discussion

This section describes and discusses the conducted experiments for performance evaluation. We have implemented an automatic framework that combines several tools and libraries. The software architecture is depicted in Figure 3. We used a Python version of 3.7 and worked on anaconda 3 with the following libraries and modules: Pandas, Gensim, NLTK, CLTK, PyArabic, PyAramorph, Stanford POS Tagger, and Tashaphyne. Pandas offer a Data Frame Object for quick and effective data handling along with integrated indexing; tools capable of reading and writing data between in-memory data structures and various formats such as text, Excel, and CSV files [74]. A python dictionary is a Python data structure that consists of a set of (key: value) pairs, where the keys are unique within one dictionary. The main functions of a dictionary are storing and extracting values using their keys [75]. We used nested dictionaries, where a collection of dictionaries is inside one single dictionary.



**Figure 3.** The tools and libraries that are used for semantic orientation evaluation.

##### 4.1. Dataset Balancing

We examine the proposed approach using the hotel reviews dataset of ElSahar and El-Beltagy [66] which was presented in Section 3.2. The dataset consists of unbalanced classes because it contains different sizes of positive, negative, and neutral reviews. At first, the neutral reviews were excluded based on the assumption that neutral texts are located close to the boundary of the binary classifier. Moreover, neutral texts are supposed to be less informative in comparison with clear positive or negative texts [76].

After that, we balanced the remaining positive and negative reviews using the under-sampling method. The objective of using under-sampling to balance the reviews is to gain a high performance of classification and to prevent the classifier from acting biased toward the majority group examples [77]. The random under-sampling is a non-heuristic method that is used to balance class sizes through the random elimination of majority class examples to make them equivalent to the smallest class size [78].

Table 10 shows the size of each class before and after class balancing. The balanced reviews dataset consists of 5294 hotel reviews (2647 positive reviews and 2647 negative

reviews). 3294 hotel reviews are used for domain ontology extraction using LDA and the manual approach. The main goal is to extract the domain ontology based on the available review's dataset. The remaining 2000 hotel reviews (1000 positive reviews and 1000 negative reviews) are used for ASA experiments, to evaluate the effectiveness of the proposed approach. The authors of [20] divided the reviews dataset into a similar approach for the same purposes.

**Table 10.** Dataset statistics before and after class balancing.

	Total Positive Reviews	Total Negative Reviews	Total Reviews
Unbalanced Reviews	10,775	2647	13,422
Balanced Reviews	2647	2647	5294
Portion1: Used for Ontology Extraction	1647	1647	3294
Portion2: Used for Sentiment Analysis Experiments	1000	1000	2000

#### 4.2. Lexicon Baseline Evaluation

The lexicon baseline approach is selected for the comparison since the lexicon baseline approach does not consider the domain concepts to identify review polarity; it simply used a sentiment lexicon to extract all the words from the review with their polarities. The ArSenL lexicon of Badaro et al. [67] is used in this experiment. Tables 11 and 12 present the confusion matrix and performance measures of the lexicon baseline approach.

**Table 11.** Confusion matrix for lexicon baseline approach.

		Predicted	
		Positive	Negative
Original	Positive	898	102
	Negative	404	596

**Table 12.** Performance evaluation of lexicon baseline approach.

	Precision	Recall	F1-Measure
Positive	68.97%	89.80%	78.01%
Negative	85.38%	59.60%	70.19%
Average	77.17%	74.70%	74.10%

The confusion matrix of the lexicon baseline approach shows that the number of correctly classified positive reviews is 898, and the number of correctly classified negative reviews is 596. The number of incorrectly classified positive reviews is 404, and the number of incorrectly classified negative reviews is 102. The overall precision of the lexicon baseline approach is 77.17% with a higher precision value for the negative reviews; the opposite is the case with the recall since the higher recall value is for the positive class with an overall recall of 74.70%. The overall f-measure value is 74.10%.

#### 4.3. Ontology Baseline Evaluation

The hotel ontology, built of 203 concepts and 6 levels, is used in this experiment as a domain concepts dictionary for features selection. The domain features are considered the best semantic features to represent each review. The hotel concepts, along with the noun POS tags, are used to identify the domain features and calculate their polarities using the N-gram around method with  $N = 4$ . 4 words before and 4 words after each domain feature are extracted and searched in the ArSenL lexicon to identify its polarity. The confusion matrix

of this approach is shown in Table 13. The number of true-positive reviews is 929, and the number of true-negative reviews is 638. The number of false-positive reviews is 362, and the number of false-negative reviews is 71. Table 14 presents performance measures of the ontology baseline approach. The overall precision is 80.96% with a higher precision value for the negative reviews. The overall recall is 78.35%, where the positive class obtained a higher recall value. The overall f-measure is 77.87% with the higher f-measure value for the positive class.

**Table 13.** Confusion matrix for ontology baseline approach.

		Predicted	
		Positive	Negative
Original	Positive	929	71
	Negative	362	638

**Table 14.** Performance evaluation of ontology baseline approach.

	Precision	Recall	F1-Measure
Positive	71.95%	92.90%	81.09%
Negative	89.98%	63.80%	74.66%
Average	80.96%	78.35%	77.87%

#### 4.4. Ontology with Level Importance Evaluation

The ontology with level importance approach is utilizing the ontology for both domain features extraction and domain feature importance identification based on their levels in the ontology tree. The hotel dictionary which was built based on the extracted ontology is used to determine the hotel features and their levels. The confusion matrix of this approach is depicted in Table 15. This approach predicted 938 and 644 reviews truly from the original one thousand positive reviews and one thousand negative reviews, respectively. The number of falsely predicted reviews from the original negative reviews is 356, and the number of falsely predicted reviews from the original positive reviews is 62.

**Table 15.** Confusion matrix for ontology with level importance approach.

		Predicted	
		Positive	Negative
Original	Positive	938	62
	Negative	356	644

Performance measures of ontology with the level importance approach are shown in Table 16. The negative class precision is 91.21% which is higher than the precision of the positive class, and the average precision of this approach is 81.84%. The positive class recall is 93.80% which is higher than the negative class recall, and the average recall for both classes is 79.1%. The average f-measure is 78.63% with a higher value for positive reviews.

**Table 16.** Performance evaluation of ontology with level importance approach.

	Precision	Recall	F1-Measure
Positive	72.48%	93.80%	81.77%
Negative	91.21%	64.40%	75.49%
Average	81.84%	79.10%	78.63%

#### 4.5. Ontology with Level and Frequency Importance Evaluation

In this experiment, we extracted the hotel features by matching the ontology concepts with the identification of their levels and their frequency importance, so the hotel concepts dictionary is used in this experiment to identify the three elements. Tables 17 and 18 present the confusion matrix and performance measures of ontology with the level and frequency importance approach.

**Table 17.** Confusion matrix for ontology with level and frequency importance approach.

		Predicted	
		Positive	Negative
Original	Positive	937	63
	Negative	353	647

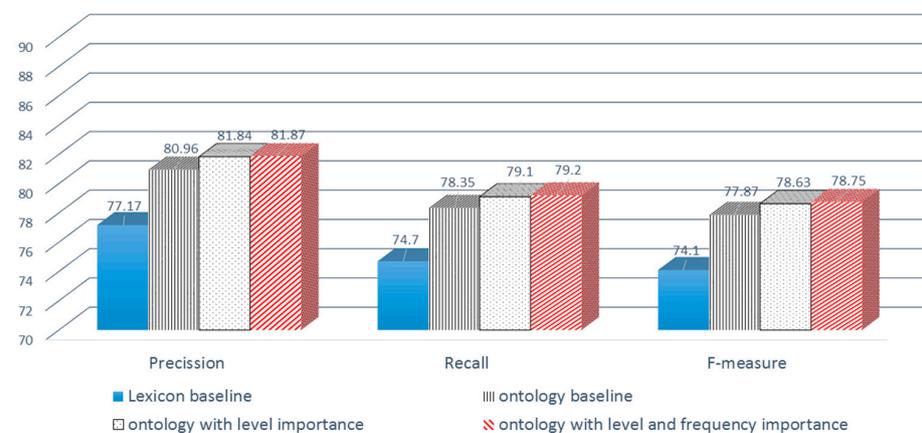
**Table 18.** Performance evaluation of ontology with level and frequency importance approach.

	Precision	Recall	F1-Measure
Positive	72.63%	93.70%	81.83%
Negative	91.12%	64.70%	75.67%
Average	81.87%	79.20%	78.75%

The number of correctly classified positive reviews using this approach is 937, and the number of correctly classified negative reviews is 647. The number of incorrectly classified positive reviews is 353, and the number of incorrectly classified negative reviews is 63. The performance measures that are presented in Table 18 demonstrated that the proposed approach achieved an overall precision of 81.87% with a higher precision value for the negative reviews, and it achieved an overall recall of 79.20% with a higher recall value for the positive class. The f-measure value is 78.75%.

#### 4.6. Results Summary and Discussion

Using ontology with domain features’ importance in the two approaches, we observed the following: the ontology with level importance and the ontology with level and frequency importance have the best results through all the semantic orientation approaches with a minor difference between them. Figure 4 summarizes results for the four schemes described earlier on average of positive and negative.



**Figure 4.** Performance evaluation of ontology for the four implemented schemes.

A comparison between different state of the art approaches for ASA is depicted in Figure 5. It reveals that the first approach yields 79.10% as accuracy. The second

approach yields 79.20% as accuracy. This may indicate that the way we utilized the concepts' frequencies in the formula needs improvement to increase the enhancement of the proposed approach. Although the difference between their performances is small, the suggested method that incorporates two factors to represent semantic domain features importance still has comparable results to other approaches.

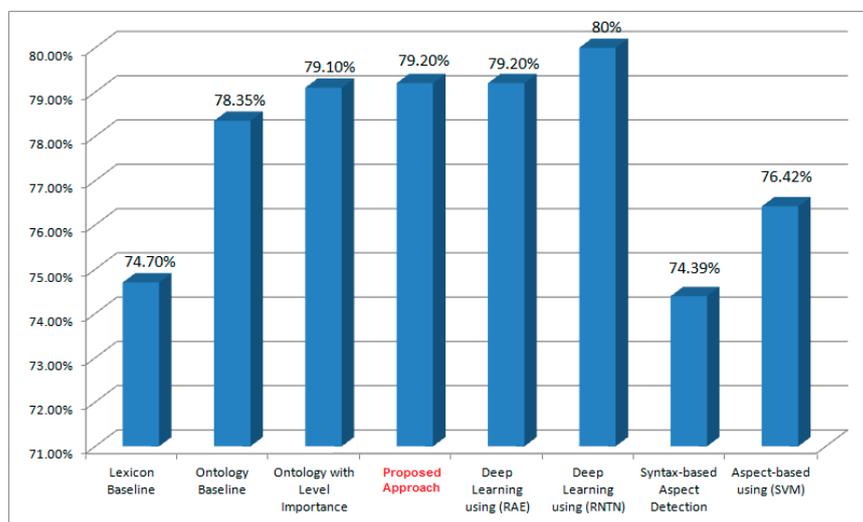


Figure 5. Accuracy of different state of the art approaches for ASA.

Combining domain ontology with the lexicon baseline approach showed an improvement up to 3.65% on accuracy value. The lexicon baseline approach did not apply any feature selection method; it just extracted all review words. Combining domain features' importance using two factors with the ontology baseline approach presents an improvement reached 0.85% for the accuracy value. Finally, the proposed approach improved the lexicon baseline approach by 4.5% for accuracy.

A comparison between the proposed approach with some state-of-the-art deep learning, machine learning, and aspect-based classifiers used for ASA is provided in Table 19. We have selected approaches that have used in common the same sentiment lexicon in [67], or the same hotels domain dataset in [66] for aspect-level-based methods.

Table 19. A comparison with some state-of-the-art approaches for ASA.

Reference	Method	Sentiment Lexicon	Dataset	Accuracy
Al-Sallab et al. [79]	Deep Learning using Recursive Auto Encoder (RAE).	ArSenL	ATB, QALB, Tweets	86.5%, 79.2%, 76.9%
Baly et al. [80]	Deep Learning using Recursive Neural Tensor Networks (RNTN).	ArSenL, ArSenTB	QALB	80%
Mataoui et al. [81]	Syntax-based Aspect Detection.	Mataoui et al. [82]	Hotels, Products	74.39%, 72.28%
Mohammad et al. [83]	Aspect-based using Support Vector Machine (SVM).	—	Hotels	76.42e%
Proposed approach	Ontology-based for Domain Features Extraction and Weighting.	ArSenL	Hotels	79.20%

Al-Sallab et al. [79] presented A Recursive Deep Learning Model for Opinion Mining in Arabic (AROMA). AROMA was tested on three Arabic datasets that were varied in writing styles and genres. Their method on the second dataset obtains an accuracy that is similar to our approach accuracy, which was (79.2%). Baly et al. [80] presented another

deep learning approach for opinion mining using Recursive Neural Tensor Networks (RNTN). Their method obtains a slightly higher accuracy rate than our approach, where the best value of accuracy was 80%. Mataoui et al. [81] and Mohammad et al. [83] methods were based on aspects of detection and extraction of hotel datasets. In comparison with their experimentation results, which were 74.39% and 76.42% accuracy, respectively, our proposed approach of sentiment analysis based on domain aspects detection, outperformed the first method accuracy by 4.81%, and the second method by 2.78%.

## 5. Conclusions

In this paper, we propose a semantic orientation approach for ASA using ontology. It incorporates a semantic domain features importance weighting method. The approach works at the feature level using an ontology of the domain concepts to extract the semantic features. It combines different factors which are: features' levels in the ontology tree, and features' frequencies in the dataset to generate overall semantic review polarity based on domain features' importance. The conducted experiment for the ontology with the level and frequency importance approach and the obtained results from this experiment demonstrated that using the frequency importance factor along with the level importance factor as an indication for the domain feature importance can increase the performance of the lexicon baseline and ontology baseline approaches with overall accuracy and f-measure values reach to 79.20% and 78.75%, respectively. The proposed approach can be comparable with the state-of-the-art methods for sentiment analysis in the Arabic language.

During this work, many limitations were faced, including the unavailability of suitable Arabic ontology for the selected domain and the unavailability of adequate lexicons for the different Arabic dialects. Future work can be derived based on these limitations: (1) Using a fully automatic approach to extract the domain ontology from the dataset available; (2) Building and using sentiment lexicon for different dialects in the Arabic language, as well as the lexicon that is used for the standard Arabic; (3) Building and using domain-specific sentiment lexicon for different domains.

**Author Contributions:** Conceptualization, S.M.K. and Q.A.A.-R.; methodology, S.M.K. and Q.A.A.-R.; software, S.M.K.; validation, S.M.K., D.M.; formal analysis, D.M.; investigation, S.M.K. and D.M.; resources, D.M.; data curation, S.M.K. and D.M.; writing—original draft preparation, S.M.K. and D.M.; writing—review and editing, S.M.K. and D.M.; visualization, S.M.K. and D.M.; funding acquisition, D.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

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

## References

1. Farha, I.A.; Magdy, W. From Arabic sentiment analysis to sarcasm detection: The arsarcasm dataset. In Proceedings of the 4th Workshop on Open-Source Arabic Corpora and Processing Tools, with a Shared Task on Offensive Language Detection, Marseille, France, 12 May 2020; pp. 32–39.
2. Alrumaih, A.; Al-Sabbagh, A.; Alsabah, R.; Kharrufa, H.; Baldwin, J. Sentiment analysis of comments in social media. *Int. J. Electr. Comput. Eng. (IJECE)* **2020**, *10*, 5917–5922. [[CrossRef](#)]
3. Al-Radaideh, Q.A.; Al-Abrat, M.A. An Arabic text categorization approach using term weighting and multiple reducts. *Soft Comput.* **2019**, *23*, 5849–5863. [[CrossRef](#)]
4. Manguri, K.H.; Ramadhan, R.N.; Amin, P.R.M. Twitter Sentiment Analysis on Worldwide COVID-19 Outbreaks. *Kurdistan J. Appl. Res.* **2020**, *2020*, 54–65. [[CrossRef](#)]
5. Boudad, N.; Faii, R.; Thami, R.; Chiheb, R. Sentiment analysis in Arabic: A review of the literature. *Ain Shams Eng. J.* **2017**, *9*, 2479–2490. [[CrossRef](#)]
6. Ghallab, A.; Mohsen, A.; Ali, Y. Arabic sentiment analysis: A systematic literature review. *Appl. Comput. Intell. Soft Comput.* **2020**, *2020*, 1–21. [[CrossRef](#)]

7. Nassr, Z.; Sael, N.; Benabbou, F. A comparative study of sentiment analysis approaches. In Proceedings of the 4th International Conference on Smart City Applications, Casablanca, Morocco, 2–4 October 2019; pp. 1–8. [\[CrossRef\]](#)
8. Atoum, J.O.; Nouman, M. Sentiment analysis of Arabic Jordanian dialect tweets. *Int. J. Adv. Comput. Sci. Appl.* **2019**, *10*, 256–262. [\[CrossRef\]](#)
9. Zahidi, Y.; El Younoussi, Y.; Azroumahli, C. Comparative Study of the Most Useful Arabic-supporting Natural Language Processing and Deep Learning Libraries. In Proceedings of the 2019 5th International Conference on Optimization and Applications (ICOA), Kenitra, Morocco, 25–26 April 2019; pp. 1–10. [\[CrossRef\]](#)
10. Mäntylä, M.V.; Graziotin, D.; Kuutila, M. The Evolution of Sentiment Analysis—A Review of Research Topics, Venues, and Top Cited Papers. *Comput. Sci. Rev.* **2018**, *27*, 16–32. [\[CrossRef\]](#)
11. Zahidi, Y.; El Younoussi, Y.; Al-Amrani, Y. Different valuable tools for Arabic sentiment analysis: A comparative evaluation. *Int. J. Electr. Comput. Eng.* **2021**, *11*, 753–762. [\[CrossRef\]](#)
12. Alharbi, A.; Taileb, M.; Kalkatawi, M. Deep learning in Arabic sentiment analysis: An overview. *J. Inf. Sci.* **2021**, *47*, 129–140. [\[CrossRef\]](#)
13. Hitzler, P.; Harmelen, S. A Reasonable Semantic Web. *Semantic Web. Interoperability Usability Appl. IOS Press J.* **2010**, *1*, 39–44. [\[CrossRef\]](#)
14. Gontier, E.M. Web Semantic and Ontology. *Adv. Internet Things* **2015**, *5*, 15. [\[CrossRef\]](#)
15. Lakshmi, S.M.; Uma, G.V. Semantic Web based e-Learning System for Sports Domain. *Int. J. Comput. Appl.* **2010**, *8*, 21–25. [\[CrossRef\]](#)
16. Man, D. Ontologies in Computer Science. *Didact. Math.* **2013**, *31*, 43–46.
17. Al Zamil, M.G.; Al-Radaideh, Q. Automatic extraction of ontological relations from Arabic text. *J. King Saud Univ.-Comput. Inf. Sci.* **2014**, *26*, 462–472. [\[CrossRef\]](#)
18. Wu, C.W.; Liu, C.L. Ontology-based Text Summarization for Business News Articles. *Comput. Appl.* **2003**, *2003*, 389–392.
19. Thakor, P.; Sasi, S. Ontology-based sentiment analysis process for social media content. *Procedia Comput. Sci.* **2015**, *53*, 199–207. [\[CrossRef\]](#)
20. Alkadri, A.M.; ElKorany, A.M. Semantic feature based Arabic opinion mining using ontology. *Int. J. Adv. Comput. Sci. Appl.* **2016**, *7*, 577–583.
21. Santosh, D.T.; Vardhan, B.V.; Ramesh, D. Extracting product features from reviews using Feature Ontology Tree applied on LDA topic clusters. In Proceedings of the 6th International Conference on Advanced Computing (IACC), Bhimavaram, India, 27–28 February 2016; pp. 163–168. [\[CrossRef\]](#)
22. Alfonso, M.M.; Sardinha, M.R. Ontology Based Aspect Level Opinion Mining. *Int. J. Eng. Sci. Res. Technol. (IJESRT)* **2016**, *5*, 797–804.
23. Zehra, S.; Wasi, S.; Jami, S.I.; Nazir, A.; Khan, A.; Waheed, N. Ontology-based Sentiment Analysis Model for Recommendation Systems. In Proceedings of the 9th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (KEOD 2017), Funchal, Portugal, 1–3 November 2017; pp. 155–160. [\[CrossRef\]](#)
24. Salas-Zárate, M.D.P.; Medina-Moreira, J.; Lagos-Ortiz, K.; Luna-Aveiga, H.; Rodriguez-Garcia, M.A.; Valencia-García, R. Sentiment analysis on tweets about diabetes: An aspect-level approach. *Comput. Math. Methods Med.* **2017**, *2017*, 1–9. [\[CrossRef\]](#)
25. El-Halees, A.; Al-Asmar, A. Ontology based Arabic opinion mining. *J. Inf. Knowl. Manag.* **2017**, *16*, 1750028. [\[CrossRef\]](#)
26. Ali, F.; Kwak, D.; Khan, P.; El-Sappagh, S.; Ali, A.; Ullah, S.; Kwak, K.S. Transportation sentiment analysis using word embedding and ontology-based topic modeling. *Knowl.-Based Syst.* **2019**, *174*, 27–42. [\[CrossRef\]](#)
27. Ravi, K.; Ravi, V. A survey on opinion mining and sentiment analysis: Tasks, approaches and applications. *Knowl.-Based Syst.* **2015**, *89*, 14–46. [\[CrossRef\]](#)
28. Amjad, A.; Qamar, U. UAMSA: Unified approach for multilingual sentiment analysis using GATE. In Proceedings of the 6th Conference on the Engineering of Computer Based Systems, Bucharest, Romania, 2–3 September 2019; pp. 1–5. [\[CrossRef\]](#)
29. Tartir, S.; Abdul-Nabi, I. Semantic sentiment analysis in Arabic social media. *J. King Saud Univ.-Comput. Inf. Sci.* **2017**, *29*, 229–233. [\[CrossRef\]](#)
30. Poria, S.; Gelbukh, A.; Hussain, A.; Howard, N.; Das, D.; Bandyopadhyay, S. Enhanced SenticNet with affective labels for concept-based opinion mining. *IEEE Intell. Syst.* **2013**, *28*, 31–38. [\[CrossRef\]](#)
31. Siddiqui, S.; Rehman, M.A.; Daudpota, S.M.; Waqas, A. Ontology Driven Feature Engineering for Opinion Mining. *IEEE Access* **2019**, *7*, 67392–67401. [\[CrossRef\]](#)
32. Oussous, A.; Benjelloun, F.Z.; Lahcen, A.A.; Belfkih, S. ASA: A framework for Arabic sentiment analysis. *J. Inf. Sci.* **2020**, *46*, 544–559. [\[CrossRef\]](#)
33. Al-Radaideh, Q. Applications of mining arabic text: A review. Recent Trends in Computational Intelligence. In *Recent Trends in Computational Intelligence*; IntechOpen: Lonon, UK, 2020; pp. 91–109. [\[CrossRef\]](#)
34. Mulki, H.; Haddad, H.; Babaoglu, I. Modern trends in Arabic sentiment analysis: A survey. *Rev. TAL* **2018**, *58*, 15–39.
35. Ihnaini, B.; Mahmuddin, M. Lexicon-Based Sentiment Analysis of Arabic Tweets: A Survey. *J. Eng. Appl. Sci.* **2018**, *13*, 7313–7322.
36. Nithish, R.; Sabarish, S.; Kishen, M.N.; Abirami, A.M.; Askarunisa, A. An Ontology based Sentiment Analysis for mobile products using tweets. In Proceedings of the 5th International Conference on Advanced Computing (ICoAC), Chennai, India, 18–20 December 2013; pp. 342–347.

37. Lazhar, F.; Yamina, T.G. Identification of opinions in Arabic texts using ontologies. In Proceedings of the Workshop on Ubiquitous Data Mining, Montpellier, France, 27 August 2012; pp. 61–64. Available online: [https://www.lirmm.fr/ecai2012/images/stories/ecai\\_doc/pdf/workshop/W3\\_procUDMECAI2012.pdf#page=67](https://www.lirmm.fr/ecai2012/images/stories/ecai_doc/pdf/workshop/W3_procUDMECAI2012.pdf#page=67) (accessed on 24 February 2019).
38. Mahyoub, F.H.; Siddiqui, M.A.; Dahab, M.Y. Building an Arabic Sentiment Lexicon Using Semi-Supervised Learning. *J. King Saud Univ.—Comput. Inf. Sci.* **2014**, *26*, 417–424. [[CrossRef](#)]
39. Soliman, T.H.; Elmasry, M.A.; Hedar, A.; Doss, M.M. Sentiment analysis of Arabic slang comments on Facebook. *Int. J. Comput. Technol.* **2014**, *12*, 3470–3478. [[CrossRef](#)]
40. Elgeldawi, E.; Sayed, A.; Galal, A.R.; Zaki, A.M. Hyperparameter Tuning for Machine Learning Algorithms Used for Arabic Sentiment Analysis. *Informatics* **2021**, *8*, 79. [[CrossRef](#)]
41. Saberi, B.; Saad, S. Sentiment analysis or opinion mining: A review. *Int. J. Adv. Sci. Eng. Inf. Technol.* **2017**, *7*, 1660–1666.
42. Sayed, A.A.; Elgeldawi, E.; Zaki, A.M.; Galal, A.R. Sentiment Analysis for Arabic Reviews using Machine Learning Classification Algorithms. In Proceedings of the 2020 International Conference on Innovative Trends in Communication and Computer Engineering (ITCE), Aswan, Egypt, 8–9 February 2020; pp. 56–63.
43. Abdullah, M.; Hadzikadicy, M.; Shaikhz, S. SEDAT: Sentiment and Emotion Detection in Arabic Text Using CNN-LSTM Deep Learning. In Proceedings of the 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, FL, USA, 17–20 December 2018; pp. 835–840.
44. Mohammed, A.; Kora, R. Deep learning approaches for Arabic sentiment analysis. *Soc. Netw. Anal. Min.* **2019**, *9*, 1–12. [[CrossRef](#)]
45. Zahidi, Y.; El Younoussi, Y.; Al-Amrani, Y. A powerful comparison of deep learning frameworks for Arabic sentiment analysis. *Int. J. Electr. Comput. Eng.* **2021**, *11*, 745–752. [[CrossRef](#)]
46. El-Affendi, M.A.; Alrajhi, K.; Hussain, A. A novel deep learning-based multilevel parallel attention neural (MPAN) model for multidomain arabic sentiment analysis. *IEEE Access* **2021**, *9*, 7508–7518. [[CrossRef](#)]
47. Khasawneh, R.T.; Wahsheh, H.A.; Alsmadi, I.M.; Al-Kabi, M.N. Arabic sentiment polarity identification using a hybrid approach. In Proceedings of the 6th International Conference on Information and Communication Systems (ICICS), Amman, Jordan, 7–9 April 2015; pp. 148–153. [[CrossRef](#)]
48. Altawaier, M.M.; Tiun, S. Comparison of machine learning approaches on arabic twitter sentiment analysis. *Int. J. Adv. Sci. Eng. Inf. Technol.* **2016**, *6*, 1067–1073. [[CrossRef](#)]
49. Al-Rubaiee, H.; Qiu, R.; Li, D. Identifying Mubasher software products through sentiment analysis of Arabic tweets. In Proceedings of the International Conference on Industrial Informatics and Computer Systems (CIICS), Sharjah, United Arab Emirates, 13–15 March 2016; pp. 1–6.
50. Alomari, K.M.; ElSherif, H.M.; Shaalan, K. Arabic tweets sentimental analysis using machine learning. In Proceedings of the International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems, Arras, France, 27–30 June 2017; pp. 602–610. [[CrossRef](#)]
51. Alayba, A.M.; Palade, V.; England, M.; Iqbal, R. Arabic language sentiment analysis on health services. In Proceedings of the 1st International Workshop on Arabic Script Analysis and Recognition (ASAR), Nancy, France, 3–5 April 2017; pp. 114–118.
52. Alowaidi, S.; Saleh, M.; Abulnaja, M. Semantic Sentiment Analysis of Arabic Texts. *Int. J. Adv. Comput. Sci. Appl.* **2017**, *8*, 256–262. [[CrossRef](#)]
53. Al-Radaideh, Q.A.; Al-Qudah, G.Y. Application of rough set-based feature selection for Arabic sentiment analysis. *Cogn. Comput.* **2017**, *9*, 436–445. [[CrossRef](#)]
54. Tripathy, A.; Agrawal, A.; Rath, S.K. Classification of sentiment reviews using n-gram machine learning approach. *Expert Syst. Appl.* **2016**, *57*, 117–126. [[CrossRef](#)]
55. Jagdale, R.S.; Shirsat, V.S.; Deshmukh, S.N. Sentiment analysis on product reviews using machine learning techniques. *Cogn. Inform. Soft Comput.* **2019**, *768*, 639–647. [[CrossRef](#)]
56. Wang, H.; Liu, L.; Song, W.; Lu, J. Feature-based sentiment analysis approach for product reviews. *J. Softw.* **2014**, *9*, 274–279. [[CrossRef](#)]
57. Catal, C.; Nangir, M. A sentiment classification model based on multiple classifiers. *Appl. Soft Comput.* **2017**, *50*, 135–141. [[CrossRef](#)]
58. Štrimaitis, R.; Stefanovič, P.; Ramanauskaitė, S.; Slotkienė, A. Financial Context News Sentiment Analysis for the Lithuanian Language. *Appl. Sci.* **2021**, *11*, 4443. [[CrossRef](#)]
59. El-Masri, M.; Altrabsheh, N.; Mansour, H.; Ramsay, A. A web-based tool for Arabic sentiment analysis. *Procedia Comput. Sci.* **2017**, *117*, 38–45. [[CrossRef](#)]
60. Lalji, T.; Deshmukh, S. Twitter sentiment analysis using hybrid approach. *Int. Res. J. Eng. Technol.* **2016**, *3*, 2887–2890.
61. El-Halees, A.M. Arabic opinion mining using combined classification approach. In Proceedings of the International Arab Conference on Information Technology ACIT, Cairo, Egypt, 8–10 December 2011; pp. 264–271. Available online: [https://www.researchgate.net/publication/228467530\\_ARABIC\\_OPINION\\_MINING\\_USING\\_COMBINED\\_CLASSIFICATION\\_APPROACH](https://www.researchgate.net/publication/228467530_ARABIC_OPINION_MINING_USING_COMBINED_CLASSIFICATION_APPROACH) (accessed on 1 February 2019).
62. Aldayel, H.K.; Azmi, A.M. Arabic tweets sentiment analysis—a hybrid scheme. *J. Inf. Sci.* **2016**, *42*, 782–797. [[CrossRef](#)]
63. Elshakankery, K.; Ahmed, M.F. HILATSA: A hybrid Incremental learning approach for Arabic tweets sentiment analysis. *Egypt. Inform. J.* **2019**, *20*, 163–171. [[CrossRef](#)]

64. Mustafa, H.H.; Mohamed, A.; Elzanfaly, D.S. An Enhanced Approach for Arabic Sentiment Analysis. *Int. J. Artif. Intell. Appl. (IJIAIA)* **2017**, *8*, 1–14. [[CrossRef](#)]
65. Gautam, G.; Yadav, D. Sentiment analysis of twitter data using machine learning approaches and semantic analysis. In Proceedings of the Seventh International Conference on Contemporary Computing (IC3), Noida, India, 7–9 August 2014; pp. 437–442. [[CrossRef](#)]
66. ElSahar, H.; El-Beltagy, S.R. Building large Arabic multi-domain resources for sentiment analysis. In Proceedings of the International Conference on Intelligent Text Processing and Computational Linguistics, Cairo, Egypt, 14–20 April 2015; Springer: Cham, Switzerland, 2015; Volume 9042, pp. 23–34. [[CrossRef](#)]
67. Badaro, G.; Baly, R.; Hajj, H.; Habash, N.; El-Hajj, W. A large scale Arabic sentiment lexicon for Arabic opinion mining. In Proceedings of the EMNLP 2014 Workshop on Arabic Natural Language Processing (ANLP), Doha, Qatar, 25 October 2014; pp. 165–173.
68. Blei, D.M.; Ng, A.Y.; Jordan, M.I. Latent Dirichlet allocation. *J. Mach. Learn. Res.* **2003**, *3*, 993–1022.
69. Zhang, Q.; Liu, S.; Gong, D.; Tu, Q. A Latent-Dirichlet-Allocation Based Extension for Domain Ontology of Enterprise’s Technological Innovation. *International Journal of Computers. Commun. Control* **2019**, *14*, 107–123.
70. Knublauch, H.; Ferguson, R.W.; Noy, N.F.; Musen, M.A. The Protégé OWL plugin: An open development environment for semantic web applications. In Proceedings of the International Semantic Web Conference; Springer: Berlin/Heidelberg, Germany, 2014; pp. 229–243.
71. Noy, N.F.; McGuinness, D.L. Ontology Development 101: A Guide to Creating Your First Ontology. 2001. 1–25. Available online: [https://protege.stanford.edu/publications/ontology\\_development/ontology101.pdf](https://protege.stanford.edu/publications/ontology_development/ontology101.pdf) (accessed on 5 May 2019).
72. Green, S.; Manning, C.D. Better Arabic parsing: Baselines, evaluations, and analysis. In Proceedings of the 23rd International Conference on Computational Linguistics, Beijing, China, 23–27 August 2010; pp. 394–402.
73. Zerrouki, T. Tashaphyne Arabic Light Stemmer and Segmentor. Available online: <https://pypi.python.org/pypi/Tashaphyne/0.2> (accessed on 27 March 2019).
74. Pandas.PyPI. Available online: <https://pypi.org/project/pandas/> (accessed on 14 February 2019).
75. Lutz, M. *Learning Python: Powerful Object-Oriented Programming*; O’Reilly Media, Inc.: Sebastopol, CA, USA, 2013.
76. Pang, B.; Lee, L. A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics, Association for Computational Linguistics, Barcelona, Spain, 21–26 July 2004; pp. 271–278.
77. Poolsawad, N.; Kambhampati, C.; Cleland, J.G.F. Balancing class for performance of classification with a clinical dataset. In Proceedings of the World Congress on Engineering, London, UK, 2–4 July 2014; Volume 1, pp. 1–6.
78. Kotsiantis, S.; Kanellopoulos, D.; Pintelas, P. Handling imbalanced datasets: A review. *GESTS Int. Trans. Comput. Sci. Eng.* **2006**, *30*, 25–36.
79. Al-Sallab, A.; Baly, R.; Hajj, H.; Shaban, K.B.; El-Hajj, W.; Badaro, G. Aroma: A recursive deep learning model for opinion mining in arabic as a low resource language. *ACM Trans. Asian Low-Resour. Lang. Inf. Processing (TALLIP)* **2017**, *16*, 1–20. [[CrossRef](#)]
80. Baly, R.; Hajj, H.; Habash, N.; Shaban, K.B.; El-Hajj, W. A sentiment treebank and morphologically enriched recursive deep models for effective sentiment analysis in arabic. *ACM Trans. Asian Low-Resour. Lang. Inf. Processing (TALLIP)* **2017**, *16*, 1–21. [[CrossRef](#)]
81. Mataoui, M.H.; Hacine, T.E.B.; Tellache, I.; Bakhtouchi, A.; Zelmati, O. A new syntax-based aspect detection approach for sentiment analysis in Arabic reviews. In Proceedings of the 2018 2nd International Conference on Natural Language and Speech Processing (ICNLSP), Algiers, Algeria, 25–26 April 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–6.
82. Mataoui, M.H.; Zelmati, O.; Boumechache, M. A proposed lexicon-based sentiment analysis approach for the vernacular Algerian Arabic. *Res. Comput. Sci.* **2016**, *110*, 55–70. [[CrossRef](#)]
83. Mohammad, A.S.; Qwasmeh, O.; Talafha, B.; Al-Ayyoub, M.; Jararweh, Y.; Benkhelifa, E. An enhanced framework for aspect-based sentiment analysis of Hotels’ reviews: Arabic reviews case study. In Proceedings of the 2016 11th International Conference for Internet Technology and Secured Transactions (ICITST), Barcelona, Spain, 5–7 December 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 98–103.