Egyptian Informatics Journal 21 (2020) 7–12

Contents lists available at ScienceDirect

Egyptian Informatics Journal

journal homepage: www.sciencedirect.com

A comprehensive study for Arabic Sentiment Analysis (Challenges and Applications)

Ahmed Alsayat [a], Nouh Elmitwally [a,b,*]

[a] Faculty of Computer and Information Sciences, Department of Computer Science, Jouf University, Saudi Arabia
[b] Faculty of Computers and Information, Department of Computer Science, Cairo University, Egypt

ARTICLE INFO

Article history:
Received 2 May 2019
Revised 21 May 2019
Accepted 19 June 2019
Available online 25 June 2019

Keywords:
Semantics
Arabic Sentiment Analysis
Social sciences
Machine learning
ALSA six-level framework

ABSTRACT

Arabic language processing works on multiple levels; less often, these complementary levels synergize well with each other. Arabic Language Sentiment Analysis (ALSA) levels consist of phonetics, morphology, syntax, lexicology, semantics, and figurative nature. The analysis of opinions and feelings is of interest in English and Indo-European languages, with little emphasis in the Arabic language, which is a language full of rhetorical characteristics and implicit meanings that have positive and negative connotations and meanings across the six linguistic levels. This paper presents a comprehensive and full proposal of a strategy for ALSA. The ALSA framework analyzes the opinions and feelings at all levels of language, in addition to the importance of building an annotated corpus, which helps to understand an Arabic sentence from the level of phonetics to the rhetorical and metonymy levels.

1. Introduction

The fields of opinion analysis, emotion development (e.g., happy, sad, and anger), and the identification of moods and emotions toward written or audio texts are vital areas in artificial intelligence applications and natural language processing. The Arabic language is unique in terms of the textual and phonetic structure of the vocabulary. Each word carries positive or negative implicit or explicit meanings. The arts and methods of linguistic texts in Arabic vary and there are many types of expression, such as prose, poetry, praise, criticism, etc. All these arts involve many semantic structures that give meanings on an explicit (verbally apparent) or implicit (rhetorical) level. Thus, a person can recognize the style of any writer of a text by looking at the meanings of the words and the feelings of positive or negative aspects of ridicule, praise, praise, booing, discontent, love-hate, etc.

Social media, such as Twitter and Facebook, has become important in understanding the attitudes of people toward a particular topic, so the data generated by social media have been used by businesses and governments to gauge opinions on specific topics or products. Many scientific studies have been conducted using data from social media to analyze users of social media in various natural languages [1].

Arabic Language Sentiment Analysis (ALSA) requires preprocessing tasks that include phonetics, morphology, sentence segmentation, part of speech tagging (POST), semantic analysis, named entities recognition (NER), subjective analysis, figurative analysis, and manual annotation for opinions using lexicons or a corpus [2–4]. Arabic sentiment analysis is typically performed using one of two basic approaches: rule-based classifiers [5,6] or machine learning classifiers, in which statistical machine learning algorithms are used to detect sentiments and opinions [7–11]. Sentiment analysis directions have been actively explored from the document level to the sentence level and the aspect level [12].

At the beginning of this study, a literary survey was conducted to identify the methods used in this type of study, which focused on the algorithms applied in Arabic. Moreover, we then identified the advances, shortcomings, and challenges when applying intelligent sciences to generate benefits at a significant and practical level. What has been achieved in such applications in the Arabic language is a minimal in comparison to the other languages. Since most of the studies were conducted in English, we tried to review

* Corresponding author.
E-mail addresses: asayat@ju.edu.sa (A. Alsayat), nselmitwally@ju.edu.sa, nouh.sabri@fci-cu.edu.eg (N. Elmitwally)
Peer review under responsibility of Faculty of Computers and Information, Cairo University.

Production and hosting by Elsevier

https://doi.org/10.1016/j.eij.2019.06.001
1110-8665/

the research work on ALSA by focusing on sentiment within the language (emotions, evaluations, feelings, opinions, sentiments, and speculations). The categorization of the sentiment part of language is one of a set of sentiment orientations (i.e., positive, negative, neutral).

The purpose of this study was to comprehensively investigate the applications and challenges of sentiment analysis of Arabic texts and how the possibility of benefiting from these applications in many areas such as figurative and rhetorical serves the precise sentiment texts in Arabic.

This research is important for the Arabic language, which suffers from a lack of interest in terms of applications of artificial intelligence, including machine learning and deep learning. According to our knowledge and as shown in the literary review, no scientific research so far has monitored and obtained findings on the processing levels of the Arabic language at a high and comprehensive level.

In this paper, we present a novel approach to Arabic sentiment analysis based on six levels of features. Our findings will help researchers in this vital research area. This paper is organized as follows: Section 2 discusses the Arabic six-level framework for the sentiment analysis process and provides a summary of the methodologies used; Sections 3 and 4 present the machine learning of ALSA and corpus annotation, respectively. In Sections 5 and 6, this paper is concluded, and future scope is provided, respectively.

2. Arabic language sentiment analysis six-level framework

2.1. Phonetic approach to Arabic sentiment analysis

No published research has involved studying the advantage of phonetic information for Arabic. We, however, think that Arabic phonetic information could be of value to the representation and sentiment analysis of the Arabic language. We think that we should write Arabic texts phonetically to know the correct pronunciation for each character or sound, considering the variety of Arabic dialect languages. Thus, along with various sentence constructions, words in Arabic can have diverse variations when written, which leads to a large number of indications.

The unique structure and vocabulary of the Arabic language affect the characteristics and attributes that distinguish it from other natural languages. The Arabic language consists of consonant and vowel characters, and each character varies depending on its pronunciation, articulation points, duration of the vowel's movement, and the following consonant characters that affect the pronunciation rules. The articulation points could be from the nose, lips, throat space, tongue, or empty mouth. The aspects of articulation point attributes could be classified as friction consonants, voice/voiceless, and soft. All this affects the sentiment meaning of the word, its impact on the mind, and how it is received in terms of positive and negative meanings [13,14].

Psychological-phonetics is the scientific study of the effect of psychological factors on the cognizant of the linguistic sounds. For example, a person who is flooded with a sense of happiness may use the word (أتراح) as joy, the word plight (محنة) as grant (منحة), and the word curse (نقمة) as blessing (نعمة). Therefore, the psychological state has a role in this psychological change of voices.

Phonetics is a science that looks at phonemes in terms of the articulation and phonology of tone and intonation. A phoneme is based on all linguistic levels: morphology, syntax, lexicon, and semantics, and indicates all characteristics (grapheme, phoneme, allophone, and morpheme).

The number of speech sounds is much higher than the number of letters and sounds in the alphabet. The articulation of phonemes/characters depends on the throat, palate, lips, and gills. The phonology of tone and intonation are voiced (المجهور), voiceless (المهموس), plosive (حرف شديد), liquid (حرف متوسط بين الشديد والرخو), fricative (الحرف الرخو الاحتكاكى), velarized (الحرف المطبق), non-velarized (الحرف المنفتح), velar (الحرف المستعلى), and non-velar (الحرف المستفل).

Ermakov et al. [15] presented a novel approach to word sentiment prediction for the Russian language based on phonetic features. Helali et al. [16] constructed a manually labelled corpus for a speech recognition model in the Arabic Tunisian language.

Our aim was to encourage researchers to pay attention to investigating Arabic phonetic characteristics in more depth based on sentiment analysis as the first level of Arabic Natural Language Processing (ANLP). The main idea involved determining the phonetic features of Arabic words that could affect their sentiment polarity.

2.2. Morphology

Morphology is a science that looks at deriving the branches of the words from their origins; the provisions of the structure of words in terms of abstraction and increase, strength, and weakness; the meaning of the morphology of derivatives; and the assets of construction and rules of analysis. Arabic is a morphologically-rich language (MRL) that interacts with sentiment analysis [17].

The most important unit of morphology, defined as having the "minimal unit of meaning", is referred to as the "morpheme". The part of speech (POS) identifies how the word is used with adjacent words that help to disambiguate the sentiment. For example, we can consider a word like "good" as an adjective that means a positive sentiment; however, it is neutral if the word is a noun. Abdul-Mageed et al. [7] showed that the context of Arabic Subjectivity and Subjective Sentiment Analysis (SSA) depends on the amount of morphological information using POS tags. Mourad et al. [18] presented a strong baseline system for performing SSA for Arabic news and tweets, depending on stemming and POS tagging. Word segmentation using POS tags has proven useful for detecting subjectivity and sentiment analysis.

Elarnaoty [19] extracted useful subjective opinions from Arabic news using a lexical parser. Moussa et al. [9] proposed a generic lexicon and developed a statistical algorithm for positive and negative sentiments based on POS tags for four categories (NOUN, VERB, ADJ, ADV).

2.3. Syntax

Syntax is a science examining the relationship of influence between the noun and verb of original sentences, and a modified sentence that achieves the purpose of matching the speech as appropriate meaning. It looks at multi-word expressions and syntactic constituents and how the arrangements of words form a sentence. Syntactic analysis extracts phrases that convey more meaning than just the individual words by themselves, such as in a noun phrase. The composition of noun phrases (NP) and verb phrases (VP) classifies the sentiment of sentences [5,20].

Ibrahim [20] presented a "sentiment analysis system for Modern Standard Arabic (MSA) and Egyptian dialect using a corpus of different types of data (tweets, product reviews, television (TV) program comments, and hotel reservations)" using syntactic features, phrase patterns, and POS n-gram tags. El-Halees [1] "used distributed representations for Arabic opinion mining and compare it with Bag of Words (BOW) representation".

2.4. Lexicology

Lexicology is the science of dictionaries that looks at all the meanings of the terms derived from the root within the tree of derivation, characterizing the Arabic language perfectly. It studies

Table 1
Arabic Language Sentiment Analysis (ALSA) levels vs. traditional sentiment/opinion levels.

Arabic Language Sentiment Analysis Levels	Traditional Sentiment/Opinion Levels	Explanation/Features
Phonetics	Aspect level	Consonants Vowels Syllables
Morphology	Aspect level	State (Indefinite, definite, construct), Gender (Masculine, feminine), Number (Singular, plural), Mood (Indicative, subjective), Voice (Active, passive), Person (1st, 2nd, 3rd), Part of speech (POS), Diacritic, Aspect [10].
Syntax	Sentence level	POS, n-grams of words, lexemes, Bag of Words (BOW), bag of lexemes. Syntactic dependency
Lexicology	Aspect + sentence + concept	- lexicon-based and learning-based approaches for concept mining from opinions - extract concept tags from visual contents and textual meta data - corpus annotation, lemmatization, POS tagging, and parsing.
Semantics	Concept level + document level	- Semantic features - Emotion (surprise, fear, disgust, sadness, happiness, anger) - positive lexicon, negative lexicon, neutral lexicon [26].
Figurative	–	Analogy, euphemism, hyperbole, context shift, false assertion, oxymoron/paradox, rhetorical question.

the level of words concerning their lexical meaning, and POS, as the lexeme, is a basic unit of lexical meaning.

Corpus linguistics impact lexicography, so that corpora can be used to reliably describe the lexicon, languages' grammar, word sense disambiguation, summarization, syntactic annotation, entity recognition, and machine translation [21]. Refaee et al. [10] presented a manually labelled annotated corpus to support sensitivity and sentiment analysis (SSA) of Arabic twitter feeds. Alayba et al. [22,23] built an Arabic sentiment corpus for health services. The collected data from Twitter were annotated manually by three annotators. The classification used deep neural networks and machine learning algorithms.

2.5. Semantics

Semantics is a science in which a particular linguistic or social context inspires a particular meaning of a particular term, demonstrating the importance of the study of words through text within a particular geographical environment and a given historical period.

The semantic level of Arabic processing involves the determination of what a sentence means by disambiguating words with multiple definitions based on the given context. This level entails the appropriate interpretation of the meaning of sentences. Sentiment is analyzed at the concept level to obtain the exact meanings in texts. The concept-based approach can detect the sentiments created by multi-word expression concepts. Semantic features focus on semantically hidden concepts extracted from texts.

Previous studies [3,11] provided a semantic sentiment analysis model of Arabic tweets to measure customer satisfaction toward telecom companies in Saudi Arabia. El-Beltagy et al. [24] presented Egyptian dialect sentiment lexicon. The sentiment lexicon including 380 words was manually constructed and used to collect more sentiment terms. Shoukry et al. [25] applied feature vectors to naïve Bayes (NB) and support vector machine (SVM) classifiers for 1000 Arabic tweets.

2.6. Arabic figurative and rhetoric

Figurative language, especially Arabic language (metaphoric and metonymy), is one of the most challenging issues faced by natural language processing (NLP). Arabic language, by nature, takes advantage of linguistic devices, such as metaphors, analogy, ambiguity, irony, sarcasm, euphemism, hyperbole, context shift, false assertions, oxymorons/paradox, and rhetorical questions, to communicate more complicated meanings [4]. We found little to no work exploring figurative Arabic sentiment detection in hyperbole, understatements, rhetorical questions, metaphors, similes, and other creative uses of language using figures of speech.

The hyperbole (صيغة المبالغة) consists of expressing feelings and opinions with an exaggerated tone. It is often used to make a strong impression or to emphasize meaning.

Euphemism (الكناية) lessens the details of an expression or an opinion that is supposedly unpleasant to soften the original statement meaning.

2.7. Arabic sentiment analysis vs. levels of opinion

The classical sentiment/opinion can be shown on four different levels: aspect level, sentence level, concept level, and document level [12]. As mentioned before, there are six levels of Arabic sentiment analysis: phonetics, morphology, syntax, lexicology, semantics, and figurative nature. Table 1 depicts the mapping of the current level of classical sentiment analysis and the proposed levels of processing of Arabic Language Sentiment Analysis (ALSA).

Aspect level: The aspect level identifies and extracts the aspects of sentiments from the text and then specifies polarity (positive, negative, and neutral).

Sentence Level: The sentence level classifies sentiment as a subjective or objective sentence and sets the polarity (positive or negative).

Conceptual Level: The conceptual level focuses on the semantic analysis of the sentence. It examines the inference of conceptual sentiment information about emotions.

Document Level: The document level provides a sentiment summary of the document as positive or negative.

The above comparison highlights the need to complete more research on the figurative level and to build a hybrid model that incorporates more than one level to reflect the proposed six-level model of ALSA (Fig. 1).

3. Machine learning approach for Arabic sentiment analysis

Machine learning methods have long been broadly used for sentiment analysis. Machine learning addresses sentiment analysis as

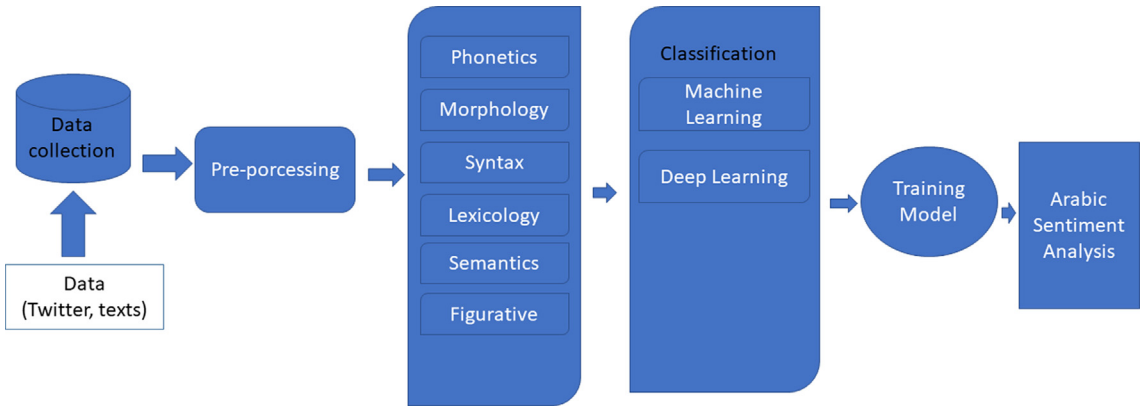

Fig. 1. Arabic Language Sentiment Analysis (ALSA) framework.

a text classification problem. Machine learning approaches include support vector machine (SVM), maximum entropy, the naïve Bayes (NB) algorithm, and artificial neural networks (ANNs). SVM has demonstrated superior performance for Arabic [11,14,19,27–29].

3.1. Sentiment classification

3.1.1. Lexicon-based methods

Sentiment lexicon of opinions and feelings involves a wide range of words and phrases that carry positive or negative meanings and are used to analyze attitudes and opinions. This lexicon is used to identify the real meanings of linguistic units [9].

3.1.2. Machine learning-based methods

We found that the most used approach is a combination of both supervised and unsupervised machine learning techniques, including SVM and K-Nearest Neighbor (KNN), in which SVM works well for a large size dataset and KNN works best for a small size corpus. However, when employing the unsupervised technique, there is no use for human-labelled data and the learning function depends on finding patterns within unlabeled data [19,27].

3.1.3. Rule-based methods

Rule-based methods primarily depend on the use of a polarity lexicon, where users set labels that are positive, negative, or neutral. Rough set theory (RST) is a mathematical tool that reduces the high dimension of the feature vector, as the main issue for sentiment analysis, to improve the classification accuracy [6].

Al-Radaideh et al. [6] proposed four different reduction computation algorithms using rough set classifiers (FRAW, Genetic, Exhaustive, Dynamic) to enhance the overall accuracy.

The Arabic original word forms and sentences can help derive a set of clear rules carrying sentiment meaning. Based on rules, assigning scores is a method of capturing the level of positive or negative sentiment within texts [5].

Table 2
A summary of the surveyed studies by Arabic-level feature.

Arabic-Level	Reference	Data Source	Sentiment Classification	Subjectivity-Type
Phonetics	–	NA	NA	NA
Morphology	[7]	DARDASHA (DAR), TAGREED (TGRD), TAHRIR (THR), MONTADA (MONT)	Support vector machine (SVM)	Subjective-positive (S-POS), subjective-negative (S-NEG)
	[19]	News corpus – manually annotated	Conditional Random Fields (CRF)	Positive, negative, or neutral orientation
	[9]	Opinosis, UMICH, UCI and ProsCons	NA	Positive, negative or neutral
	[18]	ArabSenti lexicon	Naïve Bayes (NB)	Stem POS, presence of positive emoticons
	[27]	ArabicSentiment TwitterData (ASTD)	SVM,Recursive Neural Tensor Networks (RNTN)	positive, negative, or neutral
Syntax	[20]	2000 MSA and Egyptian dialect tweets	SVM	Adjectives (JJ), nouns (NN), and verbs (VB). Positive or negative
	[1]	Hotels, books, movie reviews (OCA), Arabic Sentiment Tweets Dataset (ASTD)	SVM, Logistic Regression, and Random Forest	
Lexicology	[22]	Twitter	Machine Learning algorithms (NB, SVM, and Logistic Regression), deep and convolutional neural networks	Positive or negative
	[5]	Opinion Corpus for Arabic (OCA)	NB, SVM	Positive or negative
	[29]	Arabic Dialect tweets dataset	NB, Adaptive Boosting (AdaBoost), SVM, and Maximum Entropy (ME) algorithms	Positive or negative
Semantics	[3]	Twitter: Egyptian dialect	SVM, K-nearest neighbor (KNN), NB	Mood, opinion, attitude, and sentiment
	[24]	Twitter: Egyptian dialect	SVM	Each of the terms in the lexicon was tagged as being either a verb, adjective, noun, adverb, or idiom/compound. Positive or negative
	[2]	Twitter posts	SVM, KNN, NB	Arabic Sentiment Ontology Positive, negative, or neutral
	[25]	Twitter-Egyptian dialect	SVM and NB	Positive or negative
Figurative	–	NA	NA	NA

Thirty-two articles published from 2015 to 2019 are outlined in Table 2. The first column represents the Arabic level. The third column represents the data source and data collection. The algorithms/techniques/models that are employed and the sentiment polarity are exhibited in columns four and five, respectively.

4. Corpus annotation for Arabic language sentiment analysis

Arabic sentiment corpora, lexicons, and datasets are not often compared to other natural languages such as English. Arabic corpus annotation for sentiment analysis involves labelling features with appropriate classes in terms of meta-data for training machine learning classifiers. The Arabic annotation applies to two levels (sentence level and word level) [11,19,30]. There is a need in the Arabic annotated corpus to show an efficiently annotated corpus. The annotation can be manually performed by several native speakers, crowdsourcing, or automatically.

There is a clear need to build a high-quality annotated Arabic corpus for sentiment analysis that helps to build better classifiers and solve research issues considering the six levels of Arabic sentiment analysis. A review [31] highlighted the annotation process in recent research. Another paper [11] reported the most recent annotated corpus on Arabic Language sentiment analysis.

Al-Thubaity et al. [32] developed a corpus consisting of 5400 Saudi Dialect tweets annotated as positive, negative, neutral, objective, spam, and not sure, and with emotions (anger, disgust, fear, sadness, happiness, surprise, no emotion, and not sure). The corpus-based approach helps find sentiment words that may be positive in one field and negative in another context [12]. The research work on the topic of Arabic sentiment analysis corpus annotation has focused on the corpora type and annotation process.

5. Discussion

This research work proposes a framework for Arabic Language Sentiment Analysis (ALSA). Machine learning classifiers like NB, SVM, and decision trees are used for Arabic sentiment classification [7,18,26]. The collected datasets from social media (e.g., Twitter and Facebook) or plain texts should be manipulated to be ready for further steps. The classification process stages are: data collection, pre-process, feature selection, classification, and sentiment analysis.

As we noted from the scientific papers that focused on analysis of opinions and sentiments in Arabic, all or most of the levels of Arabic language have not been addressed in a single research work. All that was reported in the literature was that the levels are different, and each level of language was treated individually without considering the full language levels. Therefore, we urge all interested in this research area to address the subject of the analysis of opinions comprehensively and thoroughly by considering all levels of language, which may produce results that are important. The focus on Arabic language levels will lead to a significant and influential development in the subject of natural language processing in general, and automatic translation, analysis of texts, and feelings/sentiments in particular.

From the above discussion, we conclude that the ALSA framework (Fig. 1) shows promising success for comprehensive study in the area of sentiment analysis for the Arabic language (Fig. 2). There are some open issues and challenges related to Arabic sentiment analysis, including:

(1) The complexity of the Arabic language in terms of its spelling, vocabulary, phonetics, and morphology increases syntactic, semantic, and figurative ambiguity;
(2) How to translate figurative language without losing its effectual essence;
(3) The community working on Arabic Sentiment Analysis is small;
(4) Real-time sentiment analysis, spam detection, morphological faults, inadequate spelling, unstructured data, and implicit meanings;
(5) Sentiment in figurative expressions;
(6) Detecting Arabic figurative precisely for irony and sarcasm;
(7) Identifying similes, metonymy, hyperbole, and euphemism is a challenging task for humans, and it is even harder for machines;

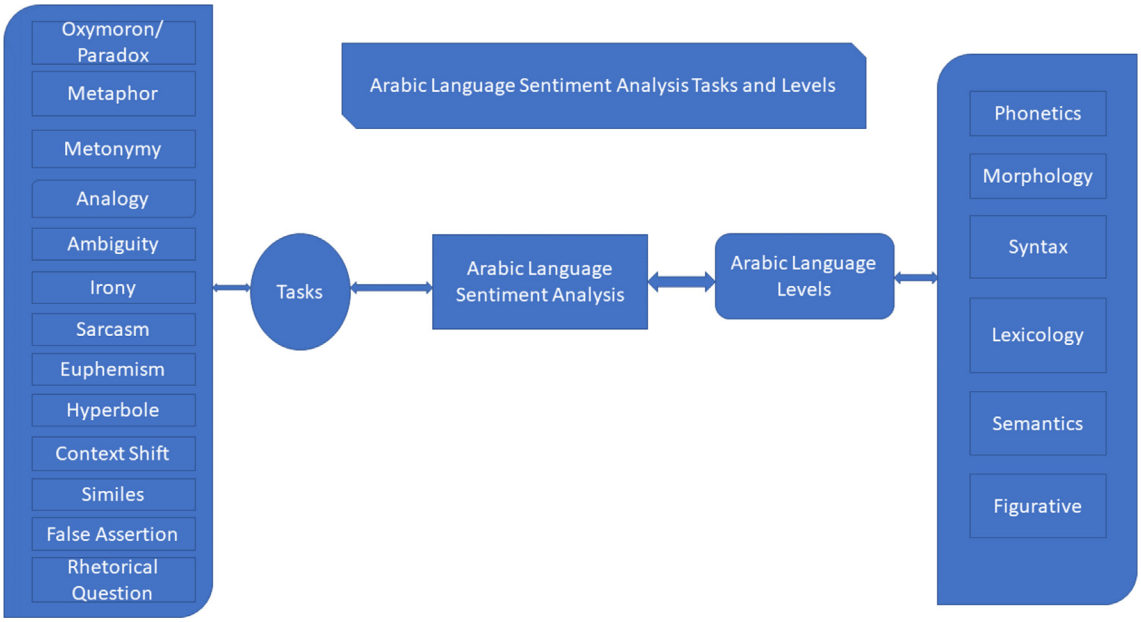

Fig. 2. ALSA tasks and levels.

(8) Synonymy and polysemy: People use the same word to refer to different meanings; and
(9) How to understand rhetorical questions.

6. Conclusions and future work

Arabic language levels can be classified by their phonetics, morphology, syntax, lexicology, semantics, and figurative nature. Arabic sentiment analysis has gained much attention in recent years. Sentiment analysis involves completely different tasks and many aspects like sentiment extraction, affective computing, subjectivity analysis, and emotional mining. Naïve Bayes and support vector machines are the most commonly used machine learning algorithms for solving the sentiment classification problem.

We have found that each level of Arabic Language Sentiment Analysis (ALSA) framework could contribute to the Arabic language sentiment analysis. However, the full hybrid model could deliver a precise sentiment meaning across the aspect, sentence and document level. In addition to that, the figurative, rhetorical (implicit meaning) would be detected ideally using the most machine and deep learning algorithms depending on a new era of building an enhanced annotated corpus that takes into its consideration the above mentioned ALSA framework.

We aim to start by building an annotated Arabic figurative corpus as future work. Then we will investigate more on the aspect level to extract the sentiment features using the machine and deep learning algorithms.

The research on Arabic Language Sentiment Analysis involves several dimensions, including figurative modeling techniques. It remains to be seen if a combination of Arabic levels will yield benefits that improve upon each level in isolation.

Funding

This research was fully funded by the Deanship of Scientific Research at Jouf University, AlJouf, Saudi Arabia on 2018/2019 with grant number (39/675) and the APC was funded by Jouf University (JU).

Declaration of Competing Interest

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

[1] El-Halees AM. Arabic opinion mining using distributed representations of documents. In: Proceedings of the 2017 Palestinian international conference on information and communication technology (PICICT), Gaza, Palestine. p. 28–33.
[2] Tartir S, Abdul-Nabi I. Semantic sentiment analysis in arabic social media. J King Saud Univ—Comput Inf Sci 2017;29:229–33.
[3] Almuqren L, Cristea AI. Framework for sentiment analysis of arabic text. In: Proceedings of the proceedings of the 27th ACM conference on hypertext and social media, Halifax, NS, Canada. p. 4.
[4] Reyes A, Rosso P. Building corpora for figurative language processing: the case of irony detection. In: Proceedings of the 4th international workshop on corpora for research on emotion sentiment & social signals, Istanbul, Turkey. p. 5.
[5] Oraby S, El-Sonbaty Y, Abou El-Nasr M. Finding opinion strength using rule-based parsing for arabic sentiment analysis. In: Castro F, Gelbukh A, González M, editors. Advances in soft computing and its applications. Berlin/Heidelberg, Germany: Springer; 2013. p. 509–20.
[6] Al-Radaideh QA, Al-Qudah GY. Application of rough set-based feature selection for Arabic sentiment analysis. Cogn. Comput. 2017;9:436–45.
[7] Abdul-Mageed M, Kuebler S, Diab M. SAMAR: a system for subjectivity and sentiment analysis of Arabic social media. Proceedings of the 3rd workshop on computational approaches to subjectivity and sentiment analysis, Jeju, Korea, 2012.
[8] Elarnaoty M, Farghaly A. Machine learning implementations in arabic text classification. In: Shaalan K, Hassanien AE, Tolba F, editors. Intelligent natural language processing: trends and applications. Cham, Switzerland: Springer International Publishing; 2018. p. 295–324.
[9] Moussa ME, Mohamed EH, Haggag MH. A generic lexicon-based framework for sentiment analysis. Int J Comput Appl 2018;1–11. doi: https://doi.org/10.1080/1206212X.2018.1483813.
[10] Refaee E, Rieser V. An Arabic twitter corpus for subjectivity and sentiment analysis. In: Proceedings of the 9th international conference on language resources and evaluation, Reykjavik, Iceland. p. 6.
[11] Almuqren L, Alzammam A, Alotaibi S, Cristea A, Alhumoud S. A review on corpus annotation for arabic sentiment analysis. In: Meiselwitz G, editor. Social computing and social media. Applications and analytics. Cham, Switzerland: Springer International Publishing; 2017. p. 215–25.
[12] Hemmatian F, Sohrabi MK. A survey on classification techniques for opinion mining and sentiment analysis. Artif Intell Rev 2017. doi: https://doi.org/10.1007/s10462-017-9599-6.
[13] Khalil A, Al-Khatib W, El-Alfy E-S, Cheded L. Anger detection in arabic speech dialogs. In: Proceedings of the 2018 international conference on computing sciences and engineering (ICCSE), Kuwait City, Kuwait. p. 1–6.
[14] Imran A, Faiyaz M, Akhtar F. An enhanced approach for quantitative prediction of personality in facebook posts. Int. J. Educ. Manag. Eng. 2018;8:8–19.
[15] Ermakov S, Ermakova L. Sentiment classification based on phonetic characteristics. In: Serdyukov P, Braslavski P, Kuznetsov SO, Kamps J, Rüger S, Agichtein E, Segalovich I, Yilmaz E, editors. Advances in information retrieval. Berlin/Heidelberg, Germany: Springer; 2013. p. 706–9.
[16] Helali W, Hajaiej Z, Cherif A. Arabic corpus implementation: application to speech recognition. In: Proceedings of the 2018 international conference on advanced systems and electric technologies (IC_ASET), Hammamet, Tunisia. p. 50–3.
[17] Ayed R, Chouigui A, Elayeb B. A new morphological annotation tool for arabic texts. In: In Proceedings of the 2018 IEEE/ACS 15th international conference on computer systems and applications (AICCSA), Aqaba, Jordan. p. 1–6.
[18] Mourad A, Darwish K. Subjectivity and sentiment analysis of modern standard arabic and arabic microblogs. Proceedings of the 4th workshop on computational approaches to subjectivity, sentiment and social media analysis, Atlanta, GA, USA, 2013.
[19] Elarnaoty M. A machine learning approach for opinion holder extraction in Arabic language. Int J Artif Intell Appl 2012;3:45–63.
[20] Ibrahim HS, Abdou SM, Gheith M. Sentiment analysis for modern standard Arabic and colloquial. Int J Nat Lang Comput 2015;4:95–109.
[21] Zeroual I, Lakhouaja A. Arabic corpus linguistics: major progress, but still a long way to go. In: Shaalan K, Hassanien AE, Tolba F, editors. Intelligent natural language processing: trends and applications. Cham, Switzerland: Springer International Publishing; 2018. p. 613–36.
[22] Alayba AM, Palade V, England M, Iqbal R. Arabic language sentiment analysis on health services. In: Proceedings of the 2017 1st international workshop on Arabic script analysis and recognition (ASAR), Nancy, France. p. 114–8.
[23] Alayba AM, Palade V, England M, Iqbal R. Improving sentiment analysis in Arabic using word representation. In: Proceedings of the 2018 IEEE 2nd international workshop on arabic and derived script analysis and recognition (ASAR), London, UK. p. 13–8.
[24] El-Beltagy SR, Ali A. Open issues in the sentiment analysis of Arabic social media: a case study. In: Proceedings of the 2013 9th international conference on innovations in information technology (IIT), Abu Dhabi, UAE. p. 215–20.
[25] Shoukry A, Rafea A. Sentence-level Arabic sentiment analysis. In: Proceedings of the 2012 international conference on collaboration technologies and systems (CTS), Denver, CO, USA. p. 546–50.
[26] Cignarella AT, Bosco C, Patti V, Lai M. Application and analysis of a multi-layered scheme for irony on the Italian Twitter Corpus TWITTIRO. In: Proceedings of the eleventh international conference on language resources and evaluation (LREC 2018), Miyazaki, Japan. p. 8.
[27] Baly R, Badaro G, El-Khoury G, Moukalled R, Aoun R, Hajj H, et al. A Characterization study of arabic twitter data with a benchmarking for state-of-the-art opinion mining models. In: Proceedings of the third Arabic natural language processing workshop, Valencia, Spain. p. 110–8.
[28] Belkebir R, Guessoum A. A Supervised approach to Arabic text summarization using AdaBoost. In: Rocha A, Correia AM, Costanzo S, Reis LP, editors. New contributions in information systems and technologies. Cham, Switzerland: Springer International Publishing; 2015. p. 227–36.
[29] Gamal D, Alfonse M, M.El-Horbaty E-S, M.Salem A-B. Twitter benchmark dataset for arabic sentiment analysis. Int J Mod. Educ Comput Sci 2019;11:33–8.
[30] Alashqar AM. A comparative study on Arabic pas tagging using Quran corpus. In: Proceedings of the 2012 8th international conference on informatics and systems (INFOS), Cairo, Egypt. p. 5.
[31] alOwisheq A, Al-Humoud S, Al-Twairesh N, AlBuhairi T. Arabic sentiment analysis resources: a survey. In: Meiselwitz G, editor. Social computing and social media. Cham, Switzerland: Springer International Publishing; 2016. p. 267–78.
[32] Al-Thubaity A, Alharbi M, Alqahtani S, Aljandal A. A Saudi dialect Twitter Corpus for sentiment and emotion analysis. In: Proceedings of the 2018 21st Saudi computer society national computer conference (NCC), Riyadh, Saudi Arabia. p. 1–6.