An Evaluation of Sentiment Analysis and Classification Algorithms for Arabic Textual Data

Ayman Mohamed Mostafa
Faculty of Computers and Informatics
Zagazig University, Egypt

ABSTRACT
Sentiment analysis is a recent advance in text mining applications for analyzing textual data according to orientation of human comments to determine whether they are positive, negative, or neutral. Different data mining techniques and algorithms such as support vector machine, naïve Bayes, decision tree, k-nearest neighbor and other techniques are used for analyzing textual data. These techniques are evaluated based on Arabic language due to its richness and diversity that can lead to difficulties in analyzing and mining large number of morphological and linguistic words that can lead to different meaning. This research provides sophisticated categorization of most or all recent articles according to the algorithms used in analyzing sentiment data. A comparison table for the proposed algorithms is presented that explains each algorithm and its use in mining and analysis of Arabic textual data and provides different evaluation for each sentiment analysis and classification algorithm according to different categories such as sentiment type, feature selection, sentiment polarity, domain oriented , data scope and data source, algorithm used in the sentiment analysis or classification, and the best algorithm result during the analysis and mining process. The experimental results explain that support vector machine algorithm presents high accuracy with approximately 77% when compared to other text mining algorithms. Different algorithms of sentiment analysis and classifications are evaluated based on their use in Arabic language which has not been evaluated before.

Keywords
Sentiment Analysis, Sentiment Classification, Arabic Textual Data, Text Mining, Support Vector Machine

1. INTRODUCTION
Sentiment Analysis (SA) or Opinion Mining (OM) is the computational study of people's opinions, attitudes and emotions toward an entity. The entity can represent individuals, events or topics that are covered by reviews. The two expressions SA or OM are interchangeable and can express a mutual meaning. Some researchers stated that OM and SA have slightly different notions [1]. Opinion mining extracts and analyzes people's opinion about an entity while sentiment analysis identifies the sentiment expressed in a text then analyzes it. Therefore, the target of SA is to find opinions, identify the sentiments they express, and then classify their polarity. There are three main classification levels in sentiment analysis (SA): document-level, sentence-level, and aspect-level SA. The document-level SA aims to classify an opinion document as expressing a positive or negative opinion or sentiment. The sentence-level SA aims to classify sentiment expressed in each sentence by identifying whether the sentence is subjective or objective. Wilson et al. [2] have pointed out that sentiment expressions are not necessarily subjective in nature. Finally, the aspect-level SA aims to classify the sentiment with respect to the specific aspects of entities.

The data sets used in SA are an important issue in this field. The main sources of data are from the product reviews. These reviews are important to the business holders as they can take business decisions according to the analysis results of users' opinions about their products. The reviews sources are mainly review sites. SA is not only applied on product reviews but can also be applied on stock markets [3, and 4], or political debates [5]. One of the recent researches that introduces a survey on different algorithms and applications for sentiment analysis is presented in [6] but it focuses on algorithms used in English, Chinese, Spanish, Italian, Dutch, and Japanese languages and did not focus on Arabic language. This paper presents a comprehensive evaluation for most recent algorithms and techniques used in opinion mining and sentiment analysis based on Arabic language.

The contribution of this research is significant for different reasons. First, this research evaluation provides sophisticated categorization of most or all recent articles according to the algorithms used. Second, various algorithms of sentiment analysis are categorized based on their use in Arabic language which has not evaluated before. Finally, the available benchmarks data sets are discussed and categorized to determine the polarity of sentiment analysis. This paper is organized as follows: section 2 includes evaluation for different Arabic sentiment analysis researches and techniques. Section 3 includes an evaluation for sentiment classification researches and techniques. Section 4 presents a comparison table for most sentiment analysis and classification algorithms based on Arabic language. Section 5 presents an evaluation results for the proposed comparison table and finally the paper is concluded in section 6.

2. ARABIC SETIMENT ANALYSIS
Arabic sentiment analysis is the process or methodology for analyzing and mining Arabic textual data to determine its behavior orientation whether it is negative, positive, or neutral according to different data sources. The analysis of Arabic sentiments can be from natural language processing, computational linguistics, or other textual data. One of the recent researches of sentiment analysis (SA) is SAMAR [7] which is a system for subjectivity and sentiment analysis (SSA) for Arabic social media genres. The objective of this research is to classify Arabic texts whether it is objective or subjective. Another subjectivity and sentiment analysis (SSA) is presented in [8] for modern standard Arabic. In this research, the task of sentence-level SSA on Modern Standard Arabic (MSA) texts is investigated from the newswire genre. The experiments showed that using morphology-based features in the proposed models improves the system's performance.

An opinion mining and analysis for Arabic language is presented in [9]. In this research, a lexicon-based tool is developed for Arabic opinion mining to analyze and process Arabic opinions collected from social media web sites. Fig. 1 presents a schematic overview of the proposed approach which consists of five phases: dataset collection, text normalization, specific features extraction from the opinions, lexicons creation, finally using classification algorithms to classify opinions into several categories based on the lexicon that was built.

The same research target is presented in [10] for analyzing sentiments in Arabic tweets. In this research a framework is developed for analyzing Arabic tweets as positive, negative, or neutral. The proposed framework contains seven steps starting from collecting training dataset and ending to verify the results using the proposed framework. Another advance in sentiment analysis is presented in [11]. In this research, an Arabic tweets and Facebook comments are analyzed using three different classifiers: Naïve Bayes, Support Vector Machine (SVM), and K-Nearest Neighbor (KNN). The experimental results stated that the Support Vector Machine algorithms achieves the highest precision while the K-Nearest Neighbor (K=10) achieves the highest recall.

The authors of [12] present an Arabic SA tool to compare three different lexicon construction techniques: Dictionary-based, Corpus based, and Integrated Lexicons. The proposed tool possesses many features such as the way negation and intensification. Another sentiment analysis methodology based on Arabic social media is presented in [13]. The authors of this research present taxonomy of sentiment analysis classification. In all testing experiments a 10-fold cross validation is used. The Naïve base classifier is used with a presence vector where the features are the words N-grams. The results showed that unigram yield the best performance and therefore it carried all the remaining experiments with unigrams.

Another sentiment analysis based on Arabic social media is presented in [14]. In this research, a case study is used to determine whether the process of establishing semantic orientation of Egyptian Arabic micro blogs is applicable. The objective of developing the sentiment lexicon is to determine semantic orientation for Arabic comments. Twitter and Arabic tweets are analyzed using opinion analysis methodology [15]. In this research, twitter sentiments are analyzed with Lexical based classification using Naive Bayes (NB) and Support Vector Machines (SVM). An analysis based on Lexicon and Corpus for Arabic sentiments is presented in [16]. This research starts by building a manually annotated dataset and then takes the reader through the detailed steps of building the lexicon.

Another sentiment analysis for Arabic social media is presented in [17]. In this research, Arabic sentiment analysis is conducted using a small dataset consisting of 1,000 Arabic reviews and comments collected from Facebook and Twitter social network websites. An Analytical Study based on Maktoob Case Study for Arabic sentiments is presented in [18]. In this research, a labeled data set is gathered for Arabic comments from a social network web site. A detailed analysis is discussed for the collected data set according to the views' length, the numbers of likes and dislikes, the polarity distribution and the languages used. A machine learning approach for opinion holder extraction in Arabic language is presented in [19]. This approach is used for extracting opinion holder in Arabic news without any lexical parsers by building a named entity feature and semantic field.

A Bilingual Experiments with an Arabic-English Corpus for Opinion Mining has been presented in [20]. An Arabic Opinion Corpus is extracted from specialized web pages and the corpus is translated into English, generating the EVOCA corpus (English Version of OCA). Different machine learning algorithms are used to classify the polarity in the corpora. The sentiment analysis results can be conducted for security areas as presented in [21]. In this research an agile sentiment analysis algorithm for social media is presented. The objective of this algorithm is to implement product reviews and analyzes the consumer comments about the presented products.

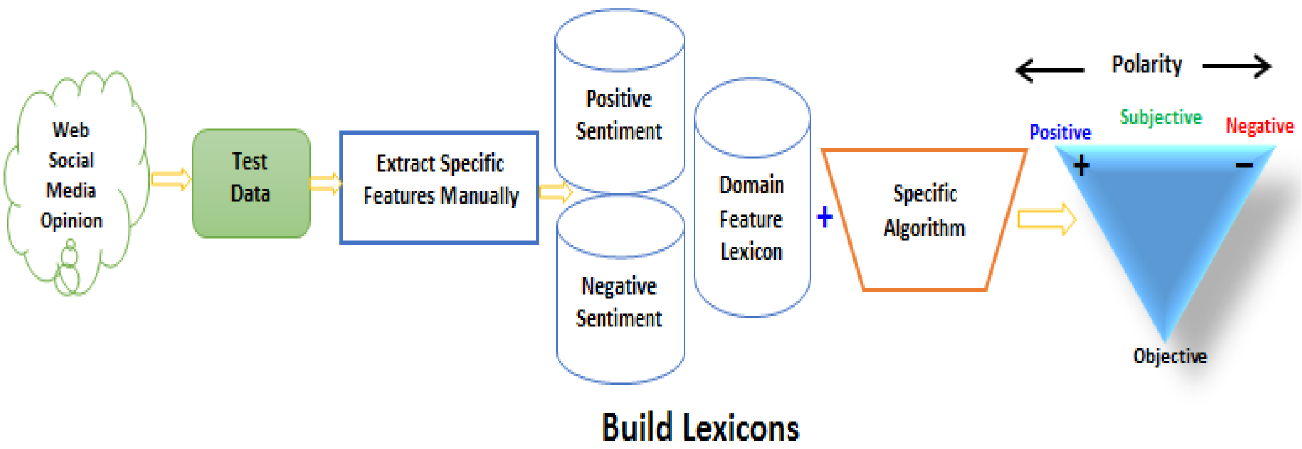

Fig. 1. Opinion analysis schema

Another sentiment analysis for Arabic business reviews is presented in [22]. The reviews are analyzed to provide their positive, negative or neutral sentiments. Based on this analysis, the users can obtain information about the local business in the language they understand, and therefore provide a better search experience for the Middle East region. One of the earlier researches in sentiment analysis is presented in [23]. This research focuses on automatic extraction of sentiment patterns from texts in Arabic language. Different sentiment words are added from recent news and the frequency of sentiment words is plotted for visualizing changes in sentiments.

3. ARABIC SETIMENT CLASSIFICATION

Arabic Sentiment classification is the process of assigning a negative, positive, or neutral classification to part of a text according to the overall opinion of the Arabic text. The sentiment classification process is considered part of the overall sentiment analysis operation. The sentiment analysis methodology starts by identifying sentiments, feature selection, sentiment classification, and finally sentiment polarity.

An Arabic sentiment lexicon is presented for assigning sentiment scores to the words found in the Arabic Word-Net [24]. A semi-supervised learning mechanism is developed by passing different positive and negative words then generate the values of the Word-Net. When the data is incomplete, uncertain or vague, a rough set theory is used for classification and analysis of data [25]. A Johnson Reducer and Genetic-based reducer algorithms are used to classify Arabic tweets without much loss of information content. . The results showed that Genetic reducer achieved high accuracy than the Johnson Reducer.

Another sentiment classification for Arabic comments on blogs, Facebook, and web sites is presented in [26]. This research is used to classify users' opinions which are performed with a new corpus for Arabic language gathered from users' posts. Another sentiment classification methodology is presented in [27]. In this research, the classification is based on short-text documentation which is considered one of the major problems in text mining. A cross lingual short-text classification is presented for Facebook comments based on Arabic and English languages. The process of topic extraction in social media is presented in [28]. The goal of the research is to develop a prototype that can determine the orientation of Arabic users with regards to a certain hot topic from Twitter.

A sentiment classification methodology for analyzing the change of Facebook status is presented in [29]. The research objective is based on focusing on the usage text mining for sentiment classification using Support Vector Machine and Naïve Bayes algorithms. A group of machine learning classifiers framework are presented in [30] for handling the problem of subjectivity and sentiment analysis for Arabic customer reviews. Three text classification algorithms: Naive Bayes, Rocchio classifier and support vector machines, are implemented as base-classifiers. Then, comparative study is made for fixed combination and meta-classifier combination.

An Arabic slang comments from social networks are presented in [31] to deal with the problems of idioms and opinion words. A sentence level sentiment analysis is also presented in [32] to deal with sentence comments instead of word comments. This can require a highly development of tools to develop the morphology and structure of Arabic language. So, the authors of [33] present a sentence classification to blogs, reviews, and tweets with regards to the given target. Different methods are proposed to address the problem of unbalanced sentiment classification in Arabic context [34]. In this research, three different methods to under-sample the majority class documents for comparing the effectiveness of the proposed methods with the random under-sampling.

An opinion mining from Arabic comparative sentences research has been presented in [35]. For identifying comparative statements from non-comparative ones, linguistic approach and machine learning approach are used. In linguistic classification, the sentence is classified based linguistic properties. A naïve search is compared against naïve Bayes for classifying sentiments in Arabic social network [36]. In this research, an application of two different approaches is presented to classify Arabic Facebook posts. The area of subjectivity and sentiment analysis (SSA) has been receiving a much interest in both the academia and the industry. As presented in [37], a Multi-Genre Corpus for Modern Standard Arabic Subjectivity and Sentiment Analysis. The unbalanced data sets are considered one of the major problems in sentiment classification. The authors of [38] address this problem by proposing three different methods to under-sample the majority class documents. These methods are: remove similar (RS), remove farthest (RF) and remove by clustering (RC).

An Arabic opinion mining methodology using combined classification approach is presented in [39]. In this research, a combined approach that automatically extracts opinions from Arabic documents is presented. A novel methodology based on intelligent agent is presented for information extraction from Arabic text without machine translation [40]. In this research, a high accuracy is achieved without the need to translate text into English language.

4. SETIMENT COMPARISON

Based on the previous detailed study on different Arabic sentiment analysis (SA) and sentiment classification (SC) algorithms, a comparison table for the proposed algorithms is presented. Table 1, explains each algorithm and its use in mining and analysis of Arabic textual data and provides different evaluation for each sentiment analysis and classification algorithm according to different categories: First: the year of publication. Second: the sentiment type whether it is sentiment analysis (SA), sentiment classification (SC), feature selection (FS) which aims to extract features from textual data, or building resource (BR) which aims to create lexica, dictionaries and corpora of Arabic data. Third: sentiment polarity whether the source of data analyzed is positive/negative (pos / neg) or general text (G). Fourth: domain oriented whether the answers during the analysis process are yes/no. Fifth: the data scope. Sixth: the data source. Sevens: the algorithm used in the sentiment analysis or classification. Eighth: the best algorithm result during the analysis and mining process.

5. EXPERIMENTAL RESULTS

Based on the sentiment analysis and classification comparison presented in Table 1, an evaluation of different algorithms used in mining and analysis of Arabic textual data is presented in Fig. 2 according to their percentage use and best performance result.

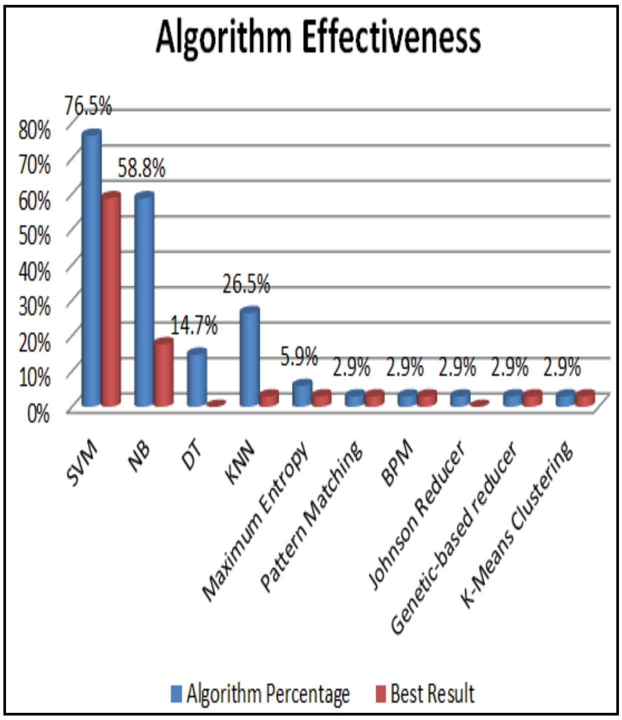

Fig. 2. Sentiment analysis algorithms' evaluation

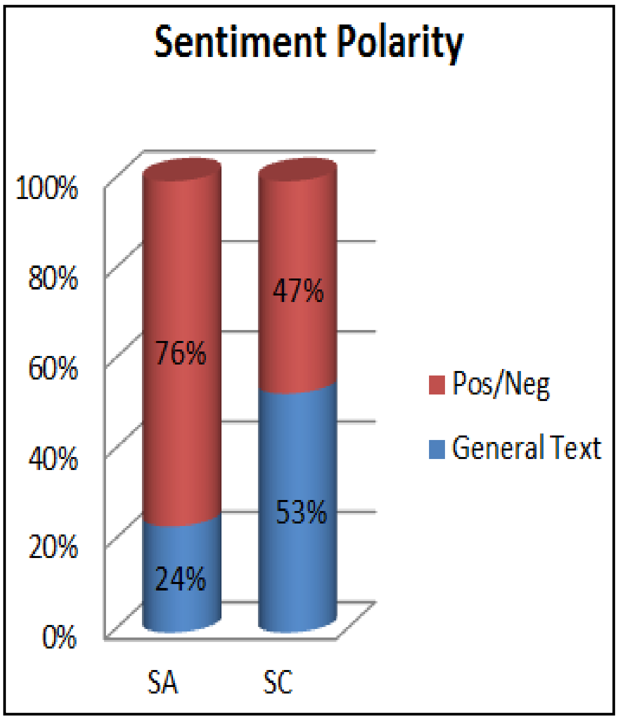

Fig. 3. Sentiment polarity

As presented in Fig. 2, support vector machine (SVM) shows efficient performance and high accuracy in many studies directed towards sentiment analysis in many languages. The SVM algorithm continues its high results in most Arabic sentiment analysis and classification datasets.

As presented in Fig. 2, approximately 77% of Arabic sentiment analysis uses support vector machine (SVM) while 59% of the applications achieve better performance using SVM. Naïve Bayes (NB) algorithm is used in 59% of sentiment analysis operations while 18% of the applications achieve better performance using NB. K-Nearest Neighbor (KNN) is used in 27% of the applications but it achieves good performance in only 3% of applications. Decision Tree (DT) is used in 15% of sentiment analysis processes but it achieves poor performance because none of the applications achieves good performance with (DT). The remaining algorithms and techniques are not used numerously because they have small effect on improving the performance of Arabic sentiment analysis and classification processes.

As presented in Fig. 3, a sentiment polarity evaluation is provided to determine the difference between (pos/neg) texts or general text (G). As shown in Fig. 3, about 76% of sentiment analysis (SA) algorithms use (pos/neg) Arabic texts while 24% of sentiment analysis algorithms use Arabic general texts (G). Approximately 47% of sentiment classification (SC) algorithms use (pos/neg) Arabic texts while 53% of sentiment classification algorithms use Arabic general texts (G). This can explain how sentiment analysis (SA) algorithms are based mainly on (pos/neg) polarity of Arabic texts because sentiment analysis analyzes textual data according to the orientation of human comments to a specific target to determine whether it is positive, negative, or neutral.

6. CONCLUSION AND FUTURE WORK

Sentiment analysis and classification of Arabic textual data are used to analyze Arabic linguistics and morphological data to determine their orientation whether the polarity of data is positive, negative, or neutral for a specific target of data. The data sets used in sentiment analysis and classification are an important issue. The main sources of data are from the product reviews. These reviews are important to the business holders can take business decisions according to the analysis results of users' opinions about their products. This research provides different categorizations of Arabic sentiment analysis in addition to an evaluation of different algorithms and techniques used in the analysis of Arabic sentiments. In future work, an algorithm for Arabic sentiment analysis will be developed based on the presented evaluation tables.

Table 1. Arabic Sentiment Analysis and Classification Comparison

Ref	Year	Sentiment Type	Sentiment Polarity	Domain Oriented	Data Scope	Data Source	Algorithm Used	Best Algorithm Result
[7]	2014	BR	Pos/Neg	N	Social Media Data	Twitter	Support Vector Machine (SVM)	Support Vector Machine (SVM)
[9]	2014	SA	Pos/Neg	N	Social Media Data	Arabic reviews social media websites	Naïve Bayes (NB) Technique, Decision Tree (DT), K-Nearest Neighbor (KNN), Support Vector Machine (SVM)	Naïve Bayes (NB)
[10]	2014	SA	G	N	Arabic tweets	Twitter	Naïve Bayes (NB), , K-Nearest Neighbor (KNN), Support Vector Machine (SVM)	Naïve Bayes (NB)
[24]	2014	SC	Pos/Neg	N	Movie Reviews	OCA Corpus Book Review	Naïve Bayes (NB), Support Vector Machine (SVM)	Naïve Bayes (NB)
[11]	2014	SA	Pos/Neg	N	Tweets and Facebook Comments	Twitter Facebook	The Naïve Bayes (NB), Support Vector Machine (SVM)and K-Nearest Neighbor (KNN)	Support Vector Machine (SVM)
[12]	2014	SA	Pos/Neg	N	Arabic Comments	Maktoob Twitter	Manual, Dictionary-based, Corpus based, and Integrated Lexicons	Integrated Lexicons
[25]	2014	SC	Pos/Neg	N	Arabic Tweets	Twitter	Johnson Reducer , Genetic-based reducer	Genetic-based reducer
[26]	2014	SC	G	N	Arabic Comments	Facebook Blogs	Support Vector Machine (SVM), Naïve Bayes (NB)	Support Vector Machine (SVM)
[27]	2014	SC	G	N	Facebook Comments	Facebook	Support Vector Machines (SVM), Naive Bayes (NB), K-Nearest Neighbor (KNN) and Decision Trees (DT)	Support Vector Machine (SVM)
[28]	2013	SC	G	N	Arabic Tweets	Twitter	K-Means Clustering Algorithm	K-Means Clustering Algorithm
[29]	2013	SC	Pos/Neg	N	Status Updates	Facebook Twitter	Support Vector Machine (SVM), Naïve Bayes (NB)	Support Vector Machine(SVM)
[13]	2013	SA	G	N	Social Media Data	Twitter	Support Vector Machine (SVM), Naïve Bayes (NB), Maximum, Bayes Net, J48 decision tree (DT).	NB better for presence vector. SVM better for frequency vector
[14]	2013	SA	Pos/Neg	N	Social Media Data	Twitter	Support Vector Machine (SVM)	Support Vector Machine (SVM)
[15]	2013	SA	Pos/Neg	Y	Social Media Data	Twitter	Naive Bayes (NB) and Support Vector Machines (SVM)	Support Vector Machines (SVM)
[30]	2013	SC	Pos/Neg	N	Arabic reviews	Www.	Naïve Bayes (NB), Rocchio	Naïve Bayes

						jeeran.com	Algorithm, Support Vector Machine (SVM), Ensemble Technique	(NB), Support Vector Machine(SVM)
[31]	2013	SC	G	N	Arabic Slang Comments	Aljazeera.net Facebook Youm7.com Alarabiya.net	Support Vector Machine (SVM)	Support Vector Machine (SVM)
[16]	2013	SA	Pos/Neg	N	Politics and Arts	Twitter	Support Vector Machine (SVM) ,Naïve Bayes (NB), Decision-Tree (DT), K-Nearest Neighbor (KNN)	Support Vector Machine (SVM)
[17]	2013	SA	Pos/Neg	N	Economic, Sport, News, Health, and Education	Facebook Twitter	Support Vector Machine (SVM) ,Naïve Bayes (NB), Decision –Tree (DT), K-Nearest Neighbor (KNN)	Support Vector Machine (SVM)

Table 1. Continue

Ref	Year	Sentiment Type	Sentiment Polarity	Domain Oriented	Data Scope	Data Source	Algorithm Used	Best Algorithm Result
[18]	2013	SA	Pos/Neg	N	Arts, Politics, Science and Technology	Yahoo Maktoob	Support Vector Machine (SVM),Naïve Bayes (NB)	Support Vector Machine (SVM)
[19]	2012	FS	Pos/Neg	N	Social Media	Twitter	Pattern Matching	Pattern Matching
[32]	2012	SC	Pos/Neg	N	Social Media Data	Twitter	Naive Bayes (NB), and Support Vector Machines (SVM)	Support Vector Machines (SVM) in Unigrams and Bigrams
[34]	2012	SC	Pos/Neg	N	Online Forums	Aljazeera	Naive Bayes (NB) Classifier, and Support Vector Machines (SVM)	Support Vector Machines (SVM)
[35]	2012	SC	G	N	Education, Technology, Sports	Corpus	Naïve Bayes (NB), K-nearest neighbors (KNN) and Support Vector Machine (SVM)	K-nearest neighbors (KNN)
[36]	2012	SC	G	N	Arabic Posts	Facebook	Naïve Bayes (NB), Naïve Search	Naïve Bayes (NB)
[37]	2012	SC	G	N	Web Forms Genres	Penn Arabic Treebank Wikipedia Talk Pages	Support Vector Machine (SVM)	Support Vector Machine (SVM)
[38]	2012	SC	Pos/Neg	N	Historical Movie Comments, Political comments	Aljazeera's web site	The Naïve Bayes (NB), Support Vector Machine (SVM) and K-Nearest Neighbor (KNN)	Support Vector Machine (SVM)
[8]	2011	SA	G	N	Corpus: College-Educated Arabic Native	Penn Arabic Treebank (PATB)	Support Vector Machine (SVM)	Support Vector Machine (SVM)

					Speakers			
[39]	2011	SC	Pos/Neg	N	Education, Politics and Sports Data	Corpus	Maximum Entropy Method, K-Nearest Neighbor (KNN), Naïve Bayes (NB), Support Vector Machine (SVM)	Maximum Entropy Method
[20]	2011	SA	Pos/Neg	N	Arabic reviews	Movie Web Pages	Support Vector Machines (SVM) and Naïve Bayes (NB)	Support Vector Machines (SVM)
[21]	2011	SA	Pos/Neg	N	Arabic Posts	Blogs	Semi-supervised (SS) sentiment classifier algorithm	Semi-supervised (SS) sentiment classifier
[33]	2010	SC	G	N	Grammatical Sentences	English Movie	Support Vector Machine (SVM)	Support Vector Machine (SVM)
[22]	2010	SA	Pos/Neg	N	Arabic Business Reviews	2,000 URLs where about 40% of them were reviews.	Machine Learning: reviews classifier	reviews classifier
[40]	2010	SC	G	N	Biographic Information	Books	Support Vector Machine (SVM), Bayes Point Machine (BPM)	Bayes Point Machine (BPM) + NLP
[23]	2005	SA	G	N	News Reports	Al Wafd Corpus	N/A	N/A

7. REFERENCES

[1] Mikalai, T., and Themis, P. "Survey on mining subjective data on the web," International Journal of Data Mining and Knowledge Discovery, vol. 24, p. 478-514, 2012

[2] Wilson, T., Wiebe, J., and Hoffman, P. "Recognizing contextual polarity in phrase-level sentiment analysis," ACM International Conference on Human Language Technology (HLT), p.347-354. 2005

[3] Hagenau, M., Liebmann, M., and Neumann, D. "Automated news reading: stock price prediction based on financial news using context-capturing features," International Journal of Decision Support Systems, ELSEVIER, vol. 55, p.685-697, 2013

[4] Liang-Chih, Y., Jheng-Long, W., Pei-Chann, C., and Hsuan-Shou, C. "Using a contextual entropy model to expand emotion words and their intensity for the sentiment classification of stock market news," International Journal of Knowledge-Based Systems, ELSEVIER, vol. 41, p. 89–97, 2013

[5] Isa, M., and Piek, V. "A lexicon model for deep sentiment analysis and opinion mining applications," International Journal of Decision Support Systems, ELSEVIER, vol. 53, p. 680–688, 2012

[6] Medhat, W., Hassan, A., and Korashy, H. "Sentiment analysis algorithms and applications: A survey," Ain Shams Engineering Journal, ELSEVIER, vol.5, p.1093-1113, 2014

[7] Abdul-Mageed, M., Diab, M., and Kubler, S. "SAMAR: Subjectivity and sentiment analysis for Arabic social media," International Journal of Computer Speech and Language, ELSEVIER, vol.28, p.20-37, 2014

[8] Abdul-Mageed, M., Diab, M. and Korayem, M. "Subjectivity and sentiment analysis of modern standard arabic," International Conference of Association for Computational Linguistics, p. 587-591, 2011

[9] Al-Kabi, M., Alsmadi, I., Gigieh, A., Wahsheh, H., and Haidar M., "Opinion Mining and Analysis for Arabic Language," International Journal of Advanced Computer Science and Applications, vol. 5, pp. 181-195, 2014

[10] Duwairi, R., Marji, R., Shaban, N., and Rushaidat, S. "Sentiment analysis in arabic tweets," IEEE International Conference on Information and Communication Systems (ICICS), p. 1-6, 2014

[11] Duwairi, R., and Qarqaz, I. "Arabic sentiment analysis using supervised classification," IEEE International Conference on Future Internet of Things and Cloud, p. 579-583, 2014

[12] Abdulla, N., Majdalawi, R., Mohammed, S., Al-Ayyoub, M., and Al-KabI, M. "Automatic Lexicon Construction for Arabic Sentiment Analysis," IEEE International Conference on Future Internet of Things and Cloud, p. 547-552, 2014

[13] Ahmed, S., Pasquier, M., and Qadah, G. "Key issues in conducting sentiment analysis on arabic social media Text," IEEE International Conference on Innovations in Information Technology (IIT),p.72-77, 2013

[14] El-Beltagy, S., and Ali, A. "Open issues in the sentiment analysis of Arabic social media: A case study," IEEE International Conference on on Innovations in Information Technology (IIT), p.215-220, 2013

[15] Ibrahim, M., and Salim, N. "Opinion analysis for twitter

and Arabic tweets: A systematic literature review," Journal of Theoretical and Applied Information Technology, vol. 56, p.338-348, 2013

[16] Abdulla, N., Ahmed, N., Shehab, M., and Al-Ayyoub, M. "Arabic sentiment analysis: lexicon-based and corpus-based," IEEE International Conference on Applied Electrical Engineering and Computing Technologies (AEECT), p.1-6, 2013

[17] Khasawneh, R., Wahsheh, H., AL-Kabi, M., and Alsmadi, I. "Sentiment analysis of arabic social media content: A comparative study," IEEE International Conference for Internet Technology and Secured Transactions (ICITST), p.101-106, 2013

[18] Al-Kabi, M., Abdulla, N., and Al-Ayyoub, M. "An analytical study of arabic sentiments: Maktoob case study," IEEE International Conference for Internet Technology and Secured Transactions (ICITST), p.89-94, 2013

[19] Elarnaoty, M., AbdelRahman, S., and Fahmy, A. "A machine learning approach for opinion holder extraction In arabic language," International Journal of Artificial Intelligence & Applications (IJAIA), vol.3, p.45-63, 2012

[20] Saleh, M., Valdivia, M., López, L., and Ortega, J. "Bilingual experiments with an arabic-english corpus for opinion mining," International Conference on Recent Advances in Natural Language Processing, p.740-745, 2011

[21] Colbaugh, R. and Glass, K. "Agile sentiment analysis of social media content for security informatics applications," IEEE International European Conference on Intelligence and Security Informatics, p.327-331, 2011

[22] Elhawary, M. and Elfeky, M. "Mining arabic business reviews," IEEE International Conference on Data Mining Workshops, p.1108-1113, 2010

[23] Ahmed, K. and Almas, Y. "Visualizing sentiments in financial texts," IEEE International Conference on Information Visualization, p.363-368, 2005

[24] Mahyoub, F., Siddiqui, M., and Dahab, M. "Building an arabic sentiment lexicon using semi-supervised learning," Journal of King Saud University–Computer and Information Sciences, vol. 26, p. 417-424, 2014

[25] Al-Radaideh, Q., and Twaiq, L. "Rough set theory for arabic sentiment classification," IEEE International Conference on Future Internet of Things and Cloud, p.559-564, 2014

[26] Akaichi, J. "Sentiment classification at the time of the Tunisian uprising," IEEE International European Conference on Network Intelligence, pp.38-45, 2014

[27] Faqeeh, M., Abdulla, N., Al-Ayyoub, Y., and Quwaider, M. "Cross-lingual short-text document classification for Facebook comments," IEEE International Conference on Future Internet of Things and Cloud, p.573-578, 2014

[28] Rafea, A., and Mostafa, N. "Topic extraction in social media," IEEE International Conference on Collaboration Technologies and Systems, p.94-98, 2013

[29] Akaichi, J., Dhouioui, Z., and Pérez, M. "Text mining Facebook status updates for sentiment classification," IEEE International Conference on System Theory, Control and Computing, p.640-645, 2013

[30] Omar, N., Albared, M., Al-Shabi, A., and Al-Moslmi, T. "Ensemble of classification algorithms for subjectivity and sentiment analysis of Arabic customers' reviews," International Journal of Advancements in Computing Technology(IJACT), vol.5, p.77-85, 2013

[31] Hassan, T., Soliman, A., and Ali, M. "Mining social networks' Arabic slang comments," International European Conference on Data Mining (ECDM), 2013

[32] Shoukry, A., and Rafea, A. "Sentence-level Arabic sentiment analysis," IEEE International Conference on Collaboration Technologies and Systems (CTS), p.546-550, 2012

[33] Farra, N., Challita, E., Abou Assi, R., and Hajj, H. "Sentence-level and document-level sentiment mining for Arabic texts", IEEE International Conference on Data Mining Workshops (ICDMW), p. 1114-1119, 2010

[34] Mountassir, A., Benbrahim, H., and Berrada, I. "Some methods to address the problem of unbalanced sentiment classification in an Arabic context," IEEE International Conference on Colloquium in Information Science and Technology (CIST), p.43-48, 2012

[35] El-Halees, A. "Opinion mining from Arabic comparative sentences," International Arab Conference on Information Technology, p.265-271, 2012

[36] Itani, M., Hamandi, L., Zantout, R., and Elkabani, I. "Classifying sentiment in Arabic social networks: Naïve search versus Naïve bayes," IEEE International Conference on Advances in Computational Tools for Engineering Applications (ACTEA), p.192-197, 2012

[37] Abdul-Mageed, M. and Diab, M. "AWATIF: A multi-genre corpus for modern standard Arabic subjectivity and sentiment analysis," International Conference on Language Resources and Evaluation (LREC), p. 3907-3914, 2012

[38] Mountassir, A., Benbrahim, H. , and Berrada, I. "An empirical study to address the problem of unbalanced data sets in sentiment classification," IEEE International Conference on Systems, Man, and Cybernetics, p.3298-3303, 2012

[39] El-Halees, A. "Arabic opinion mining using combined classification approach," International Arab Conference on Information Technology, 2011

[40] Helmy, T. and Daud, A. "Intelligent agent for information extraction from Arabic text without machine translation," International Workshop on Cross-Cultural and Cross-Lingual Aspects of the Semantic Web, 2010x