Journal of Engineering and Applied Sciences 13 (17): 7313-7322, 2018
ISSN: 1816-949X

Lexicon-Based Sentiment Analysis of Arabic Tweets: A Survey

B. Ihnaini and M. Mahmuddin
School of Computing, College of Art and Sciences, Universiti Utara Malaysia (UUM),
06010 Sintok, Kedah, Malaysia

Abstract: The quantity of data generated from Twitter and other social networks is enormous and expanding rapidly because of the growing number of users online who share their opinions and thoughts on these platforms. Extracting useful information from these data would be helpful for decision making related to services, products or people. One type of extracting information from these data is Sentiment Analysis (SA) it refers to prediction of the polarity of words to classify the expressed written feelings and opinions into positive or negative. Therefore, SA gives the organizations the ability to observe people's feelings on particular issue for example their brands and products. Although, a wide range of methods have been deployed to make such analysis but it can be used for Latin texts. On the other hand, the more complex to analyze and morphologically rich Arabic language generate a big sum of data through social media but very few analysis have been conducted on this language and its big variety of dialects. This study surveys the SA of Arabic contents, focusing on the lexicon-based methods used for extracting sentiment from Arabic Tweets written in Modern Standard Arabic (MSA) and dialectical forms. Besides, reviewing Arabic language challenges, along with going through the pre-processing tools used in the literature with some recommendations. Furthermore, showing how they generate sentiment lexicons and how they handled negation.

Key words: Sentiment analysis, lexicon-based, Twitter, modern standard Arabic

INTRODUCTION

Recently, data is considered as a currency, since, it keeps increasing in size and value. Every 2 years, the size of data is doubled (Gantz and Reinsel, 2012). The generated data in 2013 reached 4.4 Zettabytes (ZB) and it is predicted to reach 40 ZB in 2020 (Gantz and Reinsel, 2012). About 75% of these data is produced mainly by individual users (Gantz and Reinsel, 2011).

This continuous growing of data produced every second through social media, blogs and review sites is becoming larger and is hiding real value such as understanding public opinions and needs about products, policies and services. Analyzing this data could unlock the hidden value and make it useful for better and more informed decision making. SA is interested in extracting sentiment from these data to make it useful information, it aims to find the authors feeling on a certain matter in a piece of text in term of positive, neutral or negative.

There are two main approaches for SA. The first approach is supervised learning which make prediction of results based on training from annotated dataset. Data can be annotated manually by human annotators or from data itself such as number of stars obtained from the author. And to perform supervised learning approach a learning algorithm is required beside training dataset. Many earning techniques have been used such as Naive Bayes (Narayanan *et al*., 2013; Pak and Paroubek, 2010), Support Vector Machines (SVM) (Pak and Paroubek, 2010) and Decision Tree (Al-Horaibi and Khan, 2016). The second approach which his paper focuses is unsupervised called lexicon-based no training is required for this approach. It measures the polarity of a text based on the semantic orientation of the occurred words (Al-Horaibi and Khan, 2016). Therefore a lexicon containing vocabularies along with its sentiment polarity or strength is required for lexicon-based approach (Rabab'ah *et al*., 2016). Further differences between lexicon-based approach and supervised learning can be found on Table 1.

Twitter is a very attractive source for SA (Shandilya *et al*., 2017). It is the most famous micro blogging website where users can post and share messages instantly about their opinions towards variety of topics to discuss on a variety of issues and to express positive or negative feelings on products they have used

Corresponding Author: M. Mahmuddin, School of Computing, College of Art and Sciences, Universiti Utara Malaysia (UUM), 06010 Sintok, Kedah, Malaysia

Table 1: Difference between supervised learning andlexicon-based approaches

Variables	Supervised earning approach	Lexicon-based approach
Labeled dataset	Required	Not required
Training	Required	Not required
Classifier speed	Low	High
Accuracy	High	Dependent on lexical resources
Domain	Depend	Non dependent
Sentiment lexical	Optional	Required

(Agarwal *et al*., 2011). Therefore, it is rich of people's opinions. Through API provided by Twitter all Tweets are available freely (Sheela, 2016). Hence, it is the main source of data in this study.

Twitter is one of these social media platform and is considered very popular in the Arab region. The latest report of Arab statistics on social media produced by Anonymous (2017) indicates that there are more than 27 million Arabic Tweets are produced daily by 11 million Arabic users in 2017. While the researchers interest toward this language is very small. The reason for this less interest toward this language might be because of the complex structural, morphological and grammatical nature of Arabic, especially in the dialectical form (Refaee and Rieser, 2014) along with the limitation of tools and resources such as labelled corpora, dialectical lexicons and morphological analyzers (Abdul-Mageed and Diab, 2014).

MATERIALS AND METHODS

Arabic language challenges: Arabic is considered a derivational language, since, a three-letter root word can form many other words with different meanings. For example, the root (و ض ع) may form ("التواضع", "modesty", positive) or ("الوضاعة" , "rascaldom", negative). In addition, the grammatical rules of Arabic there are several forms and shapes to the same word based on its suffixes, prefixes and affixes. Often one Arabic word has more than just one affix it may be expressed as a mixture of prefixes, stems and suffixes. The prefixes are articles, conjunctions or prepositions. The suffixes are usually objects or possessive anaphora. Besides, different shapes can be written for Arabic letters. For example, the letter " ا , Alif" has four forms (ا, أ, إ, آ) (Mobarz *et al*., 2014). Therefore, this leads to some challenges to analyze Arabic text's sentiment. This highlights the big effort made with Arabic Natural Language Processing's (NLP) fundamental tools such as the morphological analyzer, syntactical parser and Part of Speech (POS) tagger.

There is still a lack of sentiment resources for the Arabic language such as annotated corpora, robust parsers and sentiment lexicons. The built resources for Arabic are not yet complete and difficult to be found by Medhat *et al.* (2014). And when compared to English, the research toward building Arabic corpora is limited (Itani *et al*., 2017). Moreover, Arabic dialect sentiment lexicons are not publically available (Abdul-Mageed and Diab, 2012).

In addition, the informality in the Arabic language on social media is considered another issue because of its big number of dialects in each country of the 22 Arab countries, besides its nature is non-grammatical and unstructured (Assiri *et al*., 2015) which cannot be analyzed by morphological analyzers tools and cannot apply POS on such language (El-Beltagy and Ali, 2013). And since, the difference between colloquial Arabic and MSA is not only the vocabulary, also, the randomness of its structure, that makes parsing this text very challenging task (Elawady *et al*., 2015).

RESULTS AND DISCUSSION

General process flow of lexicon-based SA of Arabic Tweets: SA of Arabic Tweets can be decomposed into several steps as in Fig. 1. It starts with data collection and in this study, we focus on collecting Tweets from Twitter. Therefore, most used crawling Tweets methods are reviewed in the next section. Then the preprocessing phase, where cleaning and filtering are applied on the Tweets. The next step is how researchers handled negation. Then to classify the content to positive and negative based on pre-built lexicons, the Arabic literature toward lexicons generation which is the core of the lexicon-based SA is also reviewed. The final step is to get the overall sentiment value and evaluate the performance by using a dataset collected from the first step. These stages as in Fig. 1 more described in details.

Collecting Tweets: The first step is tweet's crawling through specifying key words that gather related Tweets. Several methods are used in collecting these Tweets. First method is the Tweet crawler which gathers linked Tweets based on inquiring Twitter web service (Pak and Paroubek, 2010; Hamouda and El-Taher, 2013). The second method which is recommended in this study is to use Twitter Application Program Interface (API) which Twitter provides, this gives developers the capability of using functions like retrieving Tweets with a certain key word or a certain language using the query

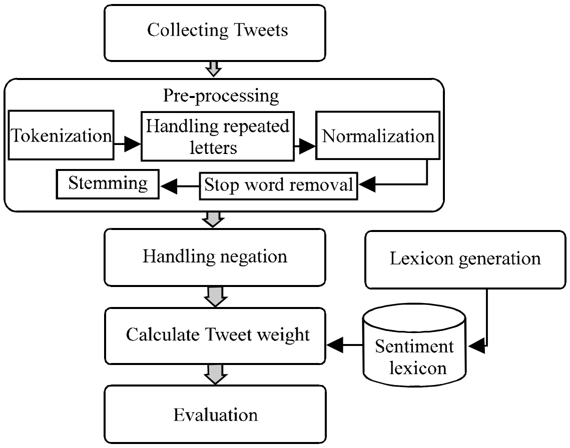

Fig. 1: General process flow of lexicon-based sentiment analysis for Arabic Tweets

"lang = ar " to retrieve content in Arabic. It is the most popular method used to collect Arabic Tweets among researchers (Al-Horaibi and Khan, 2016; Al-Horaibi and Khan, 2016; Rabab'ah *et al*., 2016); Shandilya *et al*., 2017; Agarwal *et al*., 2011; Sheela, 2016; Anonymous, 2017; El-Beltagy and Ali, 2013).

Another way to collect the Tweets is by NODEXL tool (Albraheem and Al-Khalifa, 2012). Which is an open source plug-in for Microsoft Excel and it allows automated import of any data stream from social networ servers into a spreadsheet of Excel.

Pre-processing: Pre-processing is about cleaning the unsentimental contents in a text like hashtags, user names, punctuations, URLs and pictures. Followed by tokenization stage, this technique reflects the idea of dividing the sentence to its terms. After the tokenization stage is the filtering stage which is about the correction of misspellings and phonetic errors besides removing diacritics, elongation and repeated letters is implemented.

The fast typing of users or the weak spelling skills of users leads to the misspelling. While sometimes if users wanted to express feelings about something, they repeat letters in the same word, especially vowels, like ([illegible] "alooooot "). In the classification process it is very vital to remove these repeated letters from words to identify the word. For instance, the word ([illegible] "alooooot") change to ([illegible] "a lot"), since, the vowel ([illegible]) is repeated several times. So, researchers substituted the repeated letters with only one letter. But it is suggested to tag an indicator of intensification for such words after removing the repeated letters, since, people tend to use it for intensification, for example, (greaaaaaat) is more positive than (great).

Stop words removal is applied then to those words do not have any effect on the meaning held in the text, like prepositions. That increases the speed of the analysis and the accuracy of the results. Khoja stemmer tool is an available stop word list for MSA but it includes 168 words only. And El-Khair (2006) created a list consists of 1,377 words. But the issue here is the unavailability of stop word lists that include dialectical and informal words which are used widely in social media. Therefore, it is recommended to add more stop words to the stop words lists from other dialects of Arabic, since, Tweeters do not use MSA.

Then the normalization phase where Arabic text is converted from various forms to a common form. The normalization conditions represented as following:

Remove all punctuations from the Tweets such as (. " " ;'), remove all diacritics such as ([illegible]), remove non-letters from the text such as (+ = ~ $), remove elongation (ـ) , for example, using elongation the word ([illegible]) may look like ([illegible]).

Replace ([illegible]) with bare alif (ا), replace final ([illegible]) with ([illegible]), replace final ([illegible]) with ([illegible]), replace first ([illegible]) with (ا), replace ([illegible]) with ([illegible]) and replace ([illegible]) with ([illegible]).

Finally, stemming phase where words are reduced to its base form. The obtained stem might be same as the root or different from it but it is valuable, since these words map to the same stem in general, even if the obtained stem is not a root. Table 2 highlights the main pre-processing tools used in the Arabic lexicon, based SA.

From Table 2, it have been noticed that tokenization is mandatory and it should be the first tool used for pre-processing. Then normalization and stop word removal, respectively in the same order make a must use tools for SA. And it is useful to handle the repeated letters also. Dealing with misspelling is not widely used, since, most of the informal vocabularies cannot be distinguished from misspelled words. On the same manner, part of speech tagging is also difficult to be implemented for such language. Finally, the results show that it is recommended to make stemming in the last step.

Lexicon generation: To determine polarity, data is matched with opinion words of sentiment dictionary in the lexicon-based approach. A sentiment score is given to these opinion words to show how negative, positive or objective they are in this lexicon, Identified and precompiled sentimnt terms, phrases and idioms used for communication are the sentiment lexicons that lexicon-based approaches depend on. It starts with

Table 2: Pre-processing tools for SA of Arabic text, numbering on each raw show the order of each tool

Researchers	Tokenization	Stop words	Stemming removal	Misspelling	Handling repeated Letters	Norm alization	Noun removal	Removing punctuations	Part-of-speech tagging	Others
El-Beltagy *et al.* (2013)	1	2	3							
Appel *et al.* (2016)	4		1	2		3				
Assiri *et al.* (2015)	4	5	2	1						3: Phonetic errors
Assiri *et al.* (2017)	1	5	4		2		3			
Elawady *et al.* (2015)	1	2	3							
Badaro *et al.* (2014)	1						2			
Assiri *et al.* (2017)	6	5	7		1		2		4	3:Removing diacritics
Samhaa	3			2		1				
Mataoui (2016)	1	3	2							
Mustafa *et al.* (2017)	1	4	3	2						
Al-Horaibi and Khan (2016)	1	2	5	3			4			6
Salameh *et al.* (2016)	1	2	3							
Elawady *et al.* (2015)	1	3	4	2						
Al-Harbi (2016)	1	4	2	3						
Refaee and Rieser (2016)	1	2	6	4	5	3				
Rabab'ah *et al.* (2016)	1	3	4	2						
Al-Moslmi *et al.* (2017)	2	4	3	1						
Aldayel and Azam (2015)	1	5	4	3	2					
Abd-Elhamid *et al.* (2017)	1	3	2							
Al-Kabi *et al.* (2014)	3	2	1							
Al-Kabi *et al.* (2013)	4	2	1							3:Removing non-Arabic text
Alotaibi and Khan (2017)	3	5	4	2						1:Converting emoticons to text
Total	18	17	16	4	7	15	3	4	2	
Average of order	1.6	3.4	3.9	2.5	1.8	2.1	4	2.2	4	

constructing a set of words that have sentiment orientations. Then performing classification is done through making a combination between these words. Therefore, the classification is more accurate when the used set is more comprehensive.

Many methods are implemented to generate lexicons. Some researchers built their lexicon manually (Albraheem and Al-Khalifa, 2012), some used synonyms and antonyms to expand a small set (Abdulla *et al*., 2013; Refaee and Rieser, 2014) while others used methods based on statistics and semantics such as PMI (Mohammad *et al*., 2016) and the links between the terms (Eskander and Rambow, 2015). Researches by Albraheem and Al-Khalifa (2012) have manually constructed manually a 200 word lexicon. The results have been affected by the size. They found that the stemmer is useful to reduce the size of the lexicon, since, multiple words have the same root at the lexicon. Also, Mourad and Darwish (2013) created their 3982 adjectives polarity lexicon manually. And by translating the seed list of Turney and Littman (2003) into Arabic and using it along with some random words including objective words also (Mobarz *et al*., 2014) expanded the list to obtain 150,000 words in SentiRDI.

In another research, researchers started with 300 words as a seed from SentiStrength. Then these words translated to Arabic. Two individuals annotate every word by 1 for positive words and -1 for negative words. Then using Sakhr dictionary, a synonyms list extended the lexicon. After adding emoticons to the lexicon it ended with 2376 entries (Duwairi *et al*., 2015). While ArSeLex was created manually by collectin 400 seed words then expanded manually by adding synonyms and antonyms, then automatically expanded from several online sources to form a 5244 adjectives lexicon (Ibrahim *et al*., 2015).

Building a sentiment lexicon was proposed by Al-Ayyoub *et al.* (2015) in three steps. First to collect Arabic stem, next translating these stems into English and finally. Finding the sentiment score of these stems from English sentiment lexicons located online. Around 120,000 stems have been gathered into the lexicon.

A lexicn-based method was adopted by El-Beltagy and Ali (2013) they built a lexicon form extending a seed of 380 words. Then, they annotate the lexicon's words by using two algorithms, to present the earliest lexicon with scores to its entries. Finally, they implement two approaches to find the overall sentiment of the collected Tweet's is called the sum which sums the word's polarities in a Tweet's the second is called the double polarity to find for each word in a Tweet its positive and negative score. These two scores are given to every polarity word based on the word's frequency in the corpus.

The lexicon-based approach has been under trial for improving it in Abdulla *et al*. (2014). The construction of their lexicon consisted of four stages. First, the SentiStrength website was used for selecting 300 seed words. Second, they added the synonyms of these words to the lexicon. Third, A scheme called term frequency weighting was used for identifying the missed words in the preceding steps and finally, words from different Arabic dialects were added to the lexicon for enlargement. Then a SA tool was created for calculating the text's polarity without neglecting negation and intensification. Simple lexicon-based method was used which depends on classifying the sentence based on the higher number of negative or positive words counted in it.

Three SA methods for Arabic were presented by Al-Twairesh *et al*. (2016) a simple lexicon-based method was one of them but its performance is enhanced by adding a set of features to handle valence shifters such as negation and intensification.

A comparison in performance of supervised and lexicon-based approaches have been held by Abdulla *et al*. (2013). The researchers studied corpus-based and lexicon-based approaches for Arabic SA. They developed an Arabic lexicon using a seed of 300 words and then synonyms were added to the lexicon. And aggregated all polarity weights of these words after applying negation and intensification to these weights. They indicated that the performance is poor when the lexicon is not sufficiently large in the lexicon-based approach.

SANA was presented by Abdul-Mageed and Diab (2014) as an Arabic sentiment lexicon. It combines manually created lexicons such as HUDA and SIFAAT and involves manual annotations, automatic machine translation and gloss matching by using several resources such as SAMA and THARWA. SANA includes around 225,000 entries but many of them are inflected, duplicates or not diacritized which makes it and hard to be useable. It covers two dialects, Egyptian and Levantine. And is not applied to SA tasks yet.

Badaro *et al*. (2014) built a large-scale sentiment lexicon for MSA and called it ArSenL. ArSenL consists of 28,780 lemmas and 157,969 related synsets. It is a combination of four existing resources: English WordNet (EWN), ArabicWordNet, SAMA and SentiWordNet. But it does not include dialect words but only MSA. Therefore, the accuracy is affected when it is applied on social media.

Following the example of ArSenL, Eskander and Rambow (2015) constructed Standard Arabic Sentiment Lexicon (SLSA) by developing a matching algorithm between entries in SentiWordNet and entries of an Arabic morphological analyzer. A link is create then if there is a match between the entries from both. And the score is assigned to the entries from SentiWordNet. Nevertheless, SLSA just like ArSenL has no dialect words. Hence, social medi text cannot be accurately analyzed.

Three lexicons for Arabic were generated from Twitter by Mohammad *et al*. (2016). Each one from a Twitter dataset was collected as following: the first dataset contained emoticons, the second contained a seed list of hashtags for positive and negative Arabic words and the third also contained hashtags of dialectical Arabic positive and negative words. Then three lexicons were generated using PMI from these datasets: 21, 964 for Arabic Hashtag Lexicon, 20, 128 for Dialectal Arabic Hashtag Lexicon and 43, 304 for Arabic Emoticon Lexicon.

While researches of Assiri *et al*. (2017) built an Arabic lexicon in three steps: first, from El-Beltagy and Ali (2013), they used a learning algorithm that employs seed words to expand the lexicon. In the second phase, they used a lexicon created by Badaro *et al*. (2014). This lexicon contains 154k words along with their punctuation. Since, every punctuation makes a different meaning for the words, there were different punctuations for same word in the lexicon. But the usage of punctuations by Tweeters is seldom, therefore, they removed them from the words. Then, they removed Arabic diacritics from words. Because the same word could appear in the lexicon with different diacritics, therefore, the lexicon might have duplicate text. So, they removed duplicate entries from the lexicon. In the third phase manually added new words. And they finally got a large-scale lexicon and it contains over 14k sentiment words.

Samhaa started from the earlier work of El-Beltagy and Ali (2013) and then assigned scores to the Arabic words of the lexicon in three steps. First, by collecting 100 Tweets for each word using Twitter's search API, ending with almost 500k Tweets. Second, finding co-occurrence statistics from the dataset. Finally, find the score for each word using these statistics based on the hypothesis that the stronger a polar term is, the less likely it is to co-occur with terms of an opposite polarity.

A semi-supervised approach presented by Mahyoub *et al*. (2014) to identify sentiment of Arabic text. A sentimen score was assigned for all the words of Arabic WordNet by using a small wordlist of positive and negative Arabic words as a training set first then using the relation between the words of Arabic WordNet they spread the sentiment score. The obtained lexicon consists of 800 positive words, 600 negative words and 6000 neutral words.

And in a hybrid approach proposed by Al-Moslmi *et al*. (2017) they built their own sentiment lexicon called Arabic senti-lexicon by two methods manually by three annotators and automatically based on the appearance of these words in positive/negative reviews.

Finally, Mataoui (2016) created a lexicon for the Algerian dialect only. They relied on the lexicon of El-Beltagy and Ali (2013) which its words are from Egyptian dialect and MSA. First, they remove all words not used in the Algerian dialect. After that, they added all the word of the Algerian dialect which are equivalent to Egyptian and Arabic words in the lexicon. Finally, they added the commonly used words of the Algerian dialect that carry positive or negative opinion. At the end of these steps, the constructed lexicon ended with 3093 polarity words.

Handling negation: Most research studies for Arabic SA somehow avoided handling negation (Abdulla *et al*., 2014). Negation words reverse the polarity of the sentence most of the time, changing positive sentence into negative and vice versa. Therefore, it significantly affects SA (Assiri *et al*., 2017). Some researchers mentioned that they handled negation but without showing how they did it (Abdulla *et al*., 2014; Al-Kabi *et al*., 2013). While other studies adopted the switch negation technique to handle negation which is reversing the next word to a negation word to a opposite value (Pak and Paroubek, 2010; Hamouda and El-Taher, 2013; Abdul-Mageed and Diab, 2014; Itani *et al*., 2017; Abd-Elhamid *et al*., 2016) which fails when polarity word is not next to the negation word such as in “[illegible]” where the middle word is non-polarity word means decision and the first is a negation word while the last is a polarity word means good, hence, the result should be negative in this sentence. In another simple handling of negation (Al-Horaibi and Khan, 2016) created a list of negation words that are used frequently in the Arabic language when such negation words are found in the classifier, a score of (negative = 1) is given to the word in the list. And another study multiplies the weight of the negation word by the sentiment orientation of the next word and then add it to the total sentiment orientation (Mataoui, 2016).

The negation was handled by Al-Twairesh *et al*. (2016) by checking if positive or negative word is along with the negation word. If it is positive then no increment to the counter of positive words. But if the word is negative then this time do the increment to the counter of negative words. Besides, considering the negation word’s score to be -1. Using this technique the performance increased by (1-4%) across all datasets.

Four rules to handle negation were proposed by Assiri *et al*. (2017). First to switch the word with polarity after the negation word. Second, only switch the first polarity word even if it appears two or three words after the negation word.Third if the term “[illegible]” (no) is followed by a non-polarity word first then a polarity word it makes no change. Finally, if a negation word is followed by non-polarity words then it is considered as a negative tweet such as in “not a dialogue style”. These rules increased the performance by 3%.

Evaluation: The most popular measurement to evaluate SA systems is its accuracy, therefore, a review of the accuracy of the developed Arabic lexicon-based SA have been conducted. There are two types of assigning polarity to the words in the lexicons. Either binary to have +1 or -1 as a value for words. Or it could be a score to show the weight of the word’s polarity. The average of the accuracies using the binary polarity is 66%. While using scores for polarity words achieved an average accuracy 80%. Weights were not applied to words by Albraheem and Al-Khalifa (2012) and they obtained 73% as an accuracy when tested on only 100 Tweets. Also, Abdulla *et al*. (2013) achieved an accuracy of 59.6% when tested on 2000 Tweets divided equally between Jordanian dialect and MSA. Al-Kabi *et al*. (2013) obtained an accuracy of 66% when tested on data from social media written in dialectal format and MSA. Another research using binary polarit was conducted by Abdulla *et al*. (2014) and an accuracy of 70% was achieved from using several lexicon scalability phases. While, Duwairi *et al*. (2015) obtained an accuracy of 46% when tested on 4400 Tweets. Mohammad *et al*. (2016) obtained an accuracy of 66.6% on a 300 Tweets dataset. While, Arabic SentiWordNet which have been built by Al-Horaibi and Khan (2016) achieved 60% when tested on BBN dataset, BBN. And the research of Rabab’ah *et al*. (2016) achieved an overall accuracy of 62% when they harnessed the existing SentiStrength tool for Arabic. Furthermore, Mataoui (2016) achieved 79% as the highest accuracy for binary polarity type and it was tested on 7698 comments from Facebook.

For the scored polarity type, Al-Ayyoub *et al*. (2015) achieved 86% as an accuracy when evaluated using a dataset of 900 Tweets divided equally to positive, neutral and negative Tweets. While, Eskander and Rambow (2015) tested his approach on 80 documents and achieved 68.6% as an accuracy. And an accuracy of 81% was achieved by Assiri *et al*. (2017) when tested on 4700 Tweets. Moreover an accuracy of 70% was achieved in using double polarity strategy by El-Beltagy and Ali (2013) when tested on 500 Tweets.

And a feature-based SA proposed by Abd-Elhamid *et al*. (2016) was tested on 200 MSA reviews to obtain an average accuracy of 86%. Moreover, an accuracy of 97% achieved by Al-Kabi *et al*. (2014) but the extracting and weighting features was done manually. And finally, 8 achieved an accuracy 80.3% when tested on 2000 Tweets.

CONCLUSION

In this study, a survey on all the recent lexicon-based SA for Arabic Tweets have been conducted. Focusing on how researcher's methods and techniques to generate their lexicons and how they calculate their sentiment score. And it has been found that using scores for sentiment words instead of using binary polarity words result in better accuracies with 14% difference on average. Besides, larger lexicons with higher coverage also improve the obtained accuracy.

Furthermore, it have been observed that in the pre-processing stage researchers don't use the same tools and it is recommended that tools such as tokenization, stop words removal and normalization should be used by all researchers, besides handling repeated letters instead of just removing them, since, it carries intensification of sentiment. And to finish the pre-processing phase with stemming but not aggressive stemming, since, bringing back an Arabic word to its root sometimes changes its polarity.

On the other hand, researchers did not handled negation thoroughly, since, nobody undertake which words shall be negated by the negation terms. And none of them gives a rational explanation for the effect of negation on the sentiment of words. Besides they did not consider the negation affixes such as the letter "ش" which negates the term it is attached to in most of the Arabic dialects.

Regarding the achieved results, it was discovered that despite the interesting obtained results of most of the studies but the dataset used for testing can effect on such results. Furthermore, many challenges need to be sorted out, especially, those caused by the nature of Arabic itself to get an effective and accurate SA system.

REFERENCES

Abd-Elhamid, L., D. Elzanfaly and A.S. Eldin, 2016. Feature-based sentiment analysis in online Arabic reviews. Proceedings of the 11th International Conference on Computer Engineering and Systems (ICCES'16), December 20-21, 2016, IEEE, Cairo, Egypt, ISBN:978-1-5090-3268-6, pp: 260-265.

Abdul-Mageed, M. and M.T. Diab, 2012. AWATIF: A multi-genre corpus for modern standard Arabic subjectivity and sentiment analysis. Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC'14), May 26-31, 2014, European Language Resources Association, Reykjavik, Iceland, pp: 3907-3914.

Abdul-Mageed, M. and M.T. Diab, 2014. SANA: A large scale multi-genre, multi-dialect lexicon for Arabic subjectivity and sentiment analysis. Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC'14), May 26-31, 2014, European Language Resources Association, Reykjavik, Iceland, pp: 1162-1169.

Abdulla, N., S., Mohammed, M. Al-Ayyoub and M. Al-Kabi, 2014b. Automatic lexicon construction for Arabic sentiment analysis. Proceedings of the 2014 International Conference on Future Internet of Things and Cloud (FiCloud'14), August 27-29, 2014, IEEE, Barcelona, Spain, ISBN:978-1-4799-4357-9, pp: 547-552.

Abdulla, N.A., N.A. Ahmed, M.A. Shehab and M. Al-Ayyoub, 2013. Arabic sentiment analysis: Lexicon-based and corpus-based. Proceedings of the 2013 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT'13), December 3-5, 2013, IEEE, Amman, Jordan, ISBN:978-1-4799-3676-2, pp: 1-6.

Abdulla, N.A., N.A. Ahmed, M.A. Shehab, M. Al-Ayyoub and M.N. Al-Kabi *et al*., 2014a. Towards improving the lexicon-based approach for Arabic sentiment analysis. Intl. J. Inf. Technol. Web Eng., 9: 55-71.

Agarwal, A., B. Xie, I. Vovsha, O. Rambow and R. Passonneau, 2011. Sentiment analysis of twitter data. Proceedings of the workshop on Languages in Social Media, June 23, 2011, Association for Computational Linguistics, Stroudsburg, Pennsylvania, USA., ISBN:978-1-932432-96-1, pp: 30-38.

Al-Ayyoub, M., S.B. Essa and I. Alsmadi, 2015. Lexicon-based sentiment analysis of Arabic Tweets. Intl. J. Soc. Netw. Min., 2: 101-114.

Al-Harbi, W.A., 2016. Effect of Saudi dialect preprocessing on Arabic sentiment analysis. Intl. J. Adv. Comput. Technol., 1: 91-99.

Al-Horaibi, L. and M.B. Khan , 2016. Sentiment analysis of Arabic Tweets using text mining techniques. Intl. J. Comput. Inf. Sci., 12: 100111F-100111F.

Al-Kabi, M., A. Gigieh, I. Alsmadi, H. Wahsheh and M. Haidar, 2013. An opinion analysis tool for colloquial and standard Arabic. Proceedings of the 4th International Conference on Information and Communication Systems (ICICS'13), April 23-25, 2013, ACM, Irbid, Jordan, ISBN:978-1-4503-1327-8, pp: 23-25.

Al-Kabi, M., N.M. Al-Qudah, I. Alsmadi, M. Dabour and H. Wahsheh, 2013. Arabic/English sentiment analysis: An empirical study. Proceedings of the 4th International Conference on Information and Communication Systems (ICICS'13), April 23-25, 2013, ACM, Irbid, Jordan, ISBN:978-1-4503-1327-8, pp: 23-25.

Al-Kabi, M.N., A.H. Gigieh, I.M. Alsmadi, H.A. Wahsheh and M.M. Haidar, 2014. Opinion mining and analysis for Arabic language. Intl. J. Adv. Comput. Sci. Appl., 5: 181-195.

Al-Moslmi, T., M. Albared, A. Al-Shabi, N. Omar and S. Abdullah, 2017. Arabic senti-lexicon: Constructing publicly available language resources for Arabic Sentiment analysis. J. Inf. Sci., Vol. 1,

Al-Rowaily, K., M. Abulaish, N.A.H. Haldar and M. Al-Rubaian, 2015. BiSAL-A bilingual sentiment analysis lexicon to analyze Dark Web forums for cyber security. Digital Invest., 14: 53-62.

Al-Twairesh, N., H. Al-Khalifa and A. AlSalman, 2016. AraSenTi: large-scale twitter-specific Arabic sentiment lexicons. Proceedings of the 54th Annual Meeting on Association for Computational Linguistics Vol. 1, August 7-12, 2016, Association for Computational Linguistics, Berlin, Germany, pp: 697-705.

Albraheem, L. and H.S. Al-Khalifa, 2012. Exploring the problems of sentiment analysis in informal Arabic. Proceedings of the 14th International Conference on Information Integration and Web-based Applications and Services, December 3-5, 2012, ACM, Bali, Indonesia, ISBN:978-1-4503-1306-3, pp: 415-418.

Aldayel, H.K. and A.M. Azmi, 2016. Arabic tweets sentiment analysis-a hybrid scheme. J. Inf. Sci., 42: 782-797.

Alessia, D., F. Ferri, P. Grifoni and T. Guzzo, 2015. Approaches, tools and applications for sentiment analysis implementation. Intl. J. Comput. Appl., 125: 26-33.

Alkadri, A.M. and A.M. ElKorany, 2016. Semantic feature based Arabic opinion mining using ontology. Intl. J. Adv. Comput. Sci. Appl., 7: 577-583.

Alotaibi, S. and M.B. Khan, 2017. Sentiment analysis challenges of informal Arabic language. Intl. J. Adv. Comput. Sci. Appl., 8: 278-284.

Anonymous, 2017. Social media and the internet of things. ASMedia, Taiwan.

Appel, O., F. Chiclana, J. Carter and H. Fujita, 2016. A hybrid approach to the sentiment analysis problem at the sentence level. Knowl. Based Syst., 108: 110-124.

Assiri, A., A. Emam and H. Al-Dossari, 2017. Towards enhancement of a lexicon-based approach for Saudi dialect sentiment analysis. J. Inf. Sci., 44: 184-202.

Assiri, A., A. Emam and H. Aldossari, 2015. Arabic sentiment analysis: A survey. Intl. J. Adv. Comput. Sci. Appl., 6: 75-85.

Badaro, G., R. Baly, H. Hajj, N. Habash and W. El-Hajj, 2014. A large scale Arabic sentiment lexicon for Arabic opinion mining. Proceedings of the EMNLP 2014 Workshop on Arabic Natural Language Processing (ANLP'14), October 25, 2014, Association for Computational Linguistics, Doha, Qatar, pp: 165-173.

Cozza, V. and M. Petrocchi, 2016. MIB at SemEval-2016 Task 4a: Exploiting lexicon-based features for sentiment analysis in Twitter. Proceedings of the 10th International Workshop on Semantic Evaluation, June 16-17, 2016, Association for Computational Linguistics, San Diego, California, pp: 138-143.

Duwairi, R.M., N.A. Ahmed and S.Y. Al-Rifai, 2015. Detecting sentiment embedded in Arabic social media-a lexicon-based approach. J. Intell. Fuzzy Syst., 29: 107-117.

El-Beltagy, S.R. and A. Ali, 2013. Open issues in the sentiment analysis of Arabic social media: A case study. Proceedings of the 9th International Conference on Innovations in Information Technology (IIT'13), March 17-19, 2013, IEEE, Abu Dhabi, UAE., ISBN:978-1-4673-6200-9, pp: 215-220.

El-Beltagy, S.R., 2017. WeightedNileULex: A Scored Arabic Sentiment Lexicon for Improved Sentiment Analysis. In: Language Processing, Pattern Recognition and Intelligent Systems, El-Gayar, N. and C. Suen (Eds.). World Scientific, Singapore, pp: 3-20.

El-Khair, I.A., 2006. Effects of stop words elimination for Arabic information retrieval: A comparative study. Intl. J. Comput. Inf. Sci., 4: 119-133.

Elawady, R.M., S. Barakat, H.M. El-Bakry and N. Elrashidy, 2015. Sentiment analysis for Arabic and English datasets. Intl. J. Intell. Comput. Inf. Sci., 15: 55-70.

Elhawary, M. and M. Elfeky, 2010. Mining Arabic business reviews. Proceedings of the 2010 IEEE International Conference on Data Mining Workshops (ICDMW), December 13, 2010, IEEE, Sydney, Australia, ISBN:978-1-4244-9244-2, pp: 1108-1113.

Eskander, R. and O. Rambow, 2015. SLSA: A sentiment lexicon for standard Arabic. Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, September 17-21 2015, Association for Computational Linguistics, Lisbon, Portugal, pp: 2545-2550.

Gantz, J. and D. Reinsel, 2011. Extracting value from chaos state of the universe: An executive summary. International Data Corporation Framingham, Massachusetts, USA.

Gantz, J. and D. Reinsel, 2012. The digital universe in 2020: Big data, bigger digital shadows and biggest growth in the far East. International Data Corporation Framingham, Massachusetts, USA. https://trends.ifla.org/node/88

Hamouda, A.E.D.A. and F.E.Z. El-Taher, 2013. Sentiment analyzer for arabic comments system. Intl. J. Adv. Comput. Sci. Appl., 4: 99-103.

Hemalatha, I., G.S. Varma and A. Govardhan, 2013. Sentiment analysis tool using machine learning algorithms. Intl. J. Emerging Trends Technol. Comput. Sci., 2: 105-109.

Ibrahim, H.S., S.M. Abdou and M. Gheith, 2015b. Automatic expandable large-scale sentiment lexicon of Modern Standard Arabic and Colloquial. Proceedings of the 1st International Conference on Arabic Computational Linguistics (ACLing'15), April 17-20, 2015, IEEE, Cairo, Egypt, ISBN:978-1-4673-9155-9, pp: 94-99.

Ibrahim, H.S., S.M. Abdou and M. Gheith, 2015a. Idioms-proverbs lexicon for modern standard Arabic and colloquial sentiment analysis. Intl. J. Comput. Appl., 118: 26-31.

Ibrahim, H.S., S.M. Abdou and M. Gheith, 2015. MIKA: A tagged corpus for modern standard Arabic and colloquial sentiment analysis. Proceedings of the 2015 IEEE 2nd International Conference on Recent Trends in Information Systems (ReTIS'15), July 9-11, 2015, IEEE, Kolkata, India, ISBN:978-1-4799-8349-0, pp: 353-358.

Itani, M., C. Roast and S. Al-Khayatt, 2017. Corpora for sentiment analysis of Arabic text in social media. Proceedings of the 8th International Conference on Information and Communication Systems (ICICS'17), April 4-6, 2017, IEEE, Irbid, Jordan, ISBN:978-1-5090-4244-9, pp: 64-69.

Itani, M.M., R.N. Zantout, L. Hamandi and I. Elkabani, 2012. Classifying sentiment in Arabic social networks: Naive search versus naive bayes. Proceedings of the 2nd International Conference on Advances in Computational Tools for Engineering Applications (ACTEA'12), December 12-15, 2012, IEEE, Beirut, Lebanon, ISBN:978-1-4673-2488-5, pp: 192-197.

Jurek, A., M.D. Mulvenna and Y. Bi, 2015. Improved lexicon-based sentiment analysis for social media analytics. Secur. Inf., 4: 9-21.

Khan, F.H., U. Qamar and S. Bashir, 2016. SentiMI: Introducing point-wise mutual information with SentiWordNet to improve sentiment polarity detection. Appl. Soft Comput., 39: 140-153.

Khan, F.H., U. Qamar and S. Bashir, 2017. Lexicon based semantic detection of sentiments using expected likelihood estimate smoothed odds ratio. Artif. Intell. Rev., 48: 113-138.

Kharde, V. and P. Sonawane, 2016. Sentiment analysis of twitter data: A survey of techniques. Intl. J. Comput. Appl., 139: 5-15.

Mahadzir, N.H., M.F. Omar and M.N.M. Nawi, 2016. Towards sentiment analysis application in housing projects. J. Telecommun. Electron. Comput. Eng., 8: 145-148.

Mahyoub, F.H., M.A. Siddiqui and M.Y. Dahab, 2014. Building an Arabic sentiment lexicon using semi-supervised learning. J. King Saud Univ. Comput. Inf. Sci., 26: 417-424.

Mataoui, M.H., O. Zelmati and M. Boumechache, 2016. A proposed lexicon-based sentiment analysis approach for the vernacular Algerian Arabic. Res. Comput. Sci., 110: 55-55.

Medhat, W., Hassan, A. and H. Korashy, 2014. Sentiment analysis algorithms and applications: A survey. Ain Shams Eng. J., 5: 1093-1113.

Mobarz, H., M. Rashown and I. Farag, 2014. Using automated lexical resources in Arabic sentence subjectivity. Intl. J. Artif. Intell. Appl., 5: 1-14.

Mohammad, S., M. Salameh and S. Kiritchenko, 2016. Sentiment lexicons for Arabic social media. Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC'16), May 23-28, 2016, Grand Hotel Bernardin, Portoroz, Slovenia, pp: 33-37.

Mourad, A. and K. Darwish, 2013. 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, June 14, 2013, Association for Computational Linguistics, Atlanta, Georgia, pp: 55-64.

Mustafa, M., A.S. Alsamahi and A. Hamouda, 2017. New avenues in Arabic sentiment analysis. Intl. J. Sci. Eng. Res., 8: 907-915.

Narayanan, V., I. Arora and A. Bhatia, 2013. Fast and accurate sentiment classification using an enhanced Naive Bayes model. Proceedings of the International Conference on Intelligent Data Engineering and Automated Learning, October 20-23, 2013, Springer, Berlin, Germany, pp: 194-201.

Pak, A. and P. Paroubek, 2010. Twitter as a corpus for sentiment analysis and opinion mining. Proceedings of the 7th Conference on International Language Resources and Evaluation, May 17-23, 2010, Valletta, Malta, pp: 1320-1326.

Pontiki, M., D. Galanis, J. Pavlopoulos, H. Papageorgiou and I. Androutsopoulos et al., 2014. SemEval-2014 Task 4: Aspect based sentiment analysis. Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval'14), August 23-24, 2014, Association for Computational Linguistics, Dublin, Ireland, pp: 27-35.

Rabab'ah, A.M., M. Al-Ayyoub, Y. Jararweh and M.N. Al-Kabi, 2016. Evaluating sentistrength for arabic sentiment analysis. Proceedings of the 7th International Conference on Computer Science and Information Technology (CSIT'16), July 13-14, 2016, IEEE, Amman, Jordan, ISBN:978-1-4673-8915-0, pp: 1-6.

Refaee, E. and V. Rieser, 2014. An Arabic Twitter corpus for subjectivity and sentiment analysis. Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC'14), May 26-31, 2014, European Language Resources Association, Reykjavik, Iceland, pp: 2268-2273.

Refaee, E. and V. Rieser, 2016. iLab-Edinburgh at SemEval-2016 task 7: A hybrid approach for determining sentiment intensity of Arabic Twitter phrases. Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval'16), June 16-17, 2016, Association for Computational Linguistics, San Diego, California, pp: 474-480.

Salameh, M., S. Mohammad and S. Kiritchenko, 2015. Sentiment after translation: A case-study on Arabic social media posts. Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, May 31-June 5, 2015, Association for Computational Linguistics, Denver, Colorado, pp: 767-777.

Shandilya, S.K., S. Shandilya, K. Deep and A.K. Nagar, 2017. Handbook of Research on Soft Computing and Nature-Inspired Algorithms. IGI Global, Hershey, Pennsylvania, ISBN:9781522521280, Pages: 627.

Sheela, L.J., 2016. A review of sentiment analysis in Twitter data using Hadoop. Intl. J. Database Theory Appl., 9: 77-86.

Taboada, M., J. Brooke, M. Tofiloski, K. Voll and M. Stede, 2011. Lexicon-based methods for sentiment analysis. Comput. Ling., 37: 267-307.

Tsakalidis, A., S. Papadopoulos, A.I. Cristea and Y. Kompatsiaris, 2015. Predicting elections for multiple countries using Twitter and polls. IEEE. Intell. Syst., 30: 10-17.

Turney, P.D. and M.L. Littman, 2003. Measuring praise and criticism: Inference of semantic orientation from association. Proc. ACM Trans. Inform. Syst., 21: 315-346.