Journal of Theoretical and Applied Information Technology
15th August 2017. Vol.95. No.15

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

AN AUTOMATIC LEXICON WITH EXCEPTIONAL-NEGATION ALGORITHM FOR ARABIC SENTIMENTS USING SUPERVISED CLASSIFICATION

AYMAN MOHAMED MOSTAFA
Faculty of Computers and Information, Zagazig University, Egypt
E-mail: am_mostafa@zu.edu.eg

ABSTRACT

Sentiment analysis is a kind of natural language processing that determines the feelings of people in a piece of text they are positive, negative, or neutral. Analysis of Arabic sentiments is considered a complex task due to the large linguistic and negation terms in Arabic language. Most recent researches are based on detecting the polarity term after the negation particle immediately. This can reduce the accuracy and performance of the analysis because many sentiments especially written in slang Arabic do not depend on having a negation particle before the polarity term. The aim of this paper is to develop a hybrid sentiment classification based on automatic lexicon algorithm and machine learning approach. The automatic lexicon is developed with a negation algorithm for both modern standard Arabic and colloquial Arabic. This algorithm detects the negation particle and traces all polarity terms even if they do not come after the negation particle. An exceptional negation is embedded into the negation algorithm which is based on the Arabic exceptional pattern in reversing the polarity term after the negation process. The experimental results are conducted using supervised machine learning methods such as SVM, KNN and NB that achieve high results in accuracy, precision, recall, and F-measure which are compared with the experimental results in three recent research papers.

Keywords: *Sentiment Analysis, Automatic Lexicon, Machine Learning, and Exceptional Negation*

1. INTRODUCTION

Sentiment analysis is process of analyzing reviews and comments to determine their polarity whether they are positive, negative, or neutral. Most research experiments achieved in this area have been conducted on European languages especially English and Asian languages especially Japanese and Chinese [8]. Sentiment analysis of Arabic language is more complex due to the large number of morphological and linguistics terms. Arabic language is considered highly inflectional. So, the number of studies in sentiment analysis which are conducted in Arabic whether it is expressed in modern standard Arabic (MSA) or colloquial Arabic is limited when it is compared to English sentiments [1].

One of the recent researches for studying Arabic sentiment analysis algorithms is presented in [9]. In this study, a comprehensive evaluation is presented for sentiment polarity, sentiment type, data scope, and best algorithm results in most Arabic sentiment analysis researches. A deep research of sentiment analysis challenges is presented in [19]. This research concludes that one of the major challenges of analyzing sentiments is the negation process. Different studies tried to focus on negation process on English language but in Arabic language, a small number of studies handle this problem. In addition, the mechanism for handling the polarity terms detected in negation sentences has not been presented and studied efficiently. The contribution of this paper is as follows:

- A hybrid sentiment analysis approach is presented based on lexicon-based and supervised classifications.
- Three automatic lexicons are developed for storing modern standard Arabic (MSA), colloquial Arabic, and negation terms.
- Advanced negation algorithm is implemented for increasing the flexibility of negation particles.
- Applying exceptional negation procedure for reducing false positive and false negative rates.
- High accuracy, precision, recall, and F-measure rates are achieved based on the presented intensification weights and negation algorithms.

2. RELATED WORK

The analysis process of Arabic sentiments is considered a new and a challenging domain in the last few years for different reasons. First: Arabic language contains large morphological terms. Second: it is highly inflectional language that makes the analysis of sentiments more complex. Third: Arabic language is subjective to context domains. This can make some terms may have

Journal of Theoretical and Applied Information Technology
15th August 2017. Vol.95. No.15

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

different polarities in different contexts. Fourth: colloquial Arabic in social networks is most likely to be used more than modern standard Arabic (MSA).

Other factors that make sentiment analysis difficult are that phrases can be expressed with sarcasm, irony, and/or negation [20]. A systematic review of sentiment analysis algorithms and techniques is presented in [12]. In this research, different research papers in different languages are evaluated based on their polarity and data sets. Another comparative analysis of web services is presented in [21]. The goal of this paper is to review and compare some free access web services, analyzing their capabilities to classify and score pieces of text with respect to their sentiment polarities.

An Arabic sentiment classification is presented in [13] for analyzing and mining opinions with in a new corpus. Another Arabic sentiment classification is presented in [14] for performing data reduction and minimizing redundancy using Rosetta toolkit for testing data. A short-text document classification is presented in [15] for analyzing data sets of short Facebook comments. One of the recent researches for manipulating dialectical Arabic is presented in [16]. In this research, a set of modern standard Arabic (MSA) words are converted to their equivalent meaning in dialectical words for analyzing sentiments that contain mixed expressions. A supervised classification of Arabic sentiment classification is presented in [17]. In this research, the data set is manually collected and labeled then analyzed using SVM, KNN, and NB algorithms but the precision and recall rates of the analysis are low due to the lack of data preprocessing such as stop word removal and the lack of Arabic term polarity intensification. A recent research for creating tagged corpus for MSA and colloquial Arabic terms is presented in [18] into which data sets are collected manually and a stretching process is performed by removing repeated character patterns in the same word.

3. DATA PREPROCESSING

3.1 Stemming Process

In the preprocessing of data, all comments are added according to their own lexicon. The classical comments and words are added to the modern standard Arabic (MSA) lexicon. The slang comments and terms are added to the colloquial lexicon. To increase the performance of analysis algorithms, the stemming process is used to remove all prefixes and suffixes of the word to return the word to its main root.

By collecting comments from different sources, a large number of words have been added to the lexicon with the ability to add their root in the same lexicon for increasing the performance of sentiment analysis. For example, the comment "هذا الموبايل رائع" means "This mobile is wonderful". The polarity of this comment after the analysis process is positive. The stemming process of the word "رائع" "wonderful" can generate two opposite words which are "روع" means "scare" and "مروع" means "terrible" or "fearful". These words indicate a negative polarity. These words are added to the MSA lexicon for future sentiment analysis.

3.2 Removing Stop Words

Removing stop words is used to eliminate insignificant words from the text and to minimize the size and dimension of dataset. All pronouns are removed like "أنا – نحن – أنت –هو – هى – هم" which means "I – we – you – he – she – they" respectively. Conjunction words like "أو - و" which means "or - and" respectively, are removed. Adverb words like "ثم" which means "then" is removed. Preposition letters like "من – على - فى" which means "in – on - of" respectively are removed.

3.3 Irony Detection

Irony comments are used to ridicule from specific review or subject. For example, the comment "يا سلام على النباهه" means "Oh!!! That's gumption" refers to a positive polarity but actually, it refers to a negative one. In this case, the sentiment polarity is determined according to the context of the comment meaning. If the general topic of the review is positive, then this comment will be negative polarity and vice versa.

4. AUTOMATIC LEXICON CONSTRUCTION

The process of lexicon construction is based on building three lexicons: modern standard Arabic (MSA), colloquial, and negation lexicon. These lexicons are presented as follows:

4.1 MSA and Colloquial Lexicons

The modern standard Arabic (MSA) lexicon and colloquial lexicon are used to store all classical and slang terms respectively along with their polarity weight. The general polarity weight is classified as; 3-weight, 5-weight, and 7-weight. For 3-weight, the polarity will be discrete values from 1 to -1 (1 Positive, 0 Neutral, and -1 Negative). For 5-weight, the polarity will be discrete values from 2 to -2 (2 Positive, 1 Nearly Positive, 0 Neutral, -1 Nearly Negative, and -2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Negative). For 7-weight, the polarity will be discrete values from 3 to -3 (3 High Positive, 2 Positive, 1 Nearly Positive, 0 Neutral, -1 Nearly Negative, -2 Negative, and -3 High Negative).

For lengthy sentiments and comments, more polarity terms are found. So, it is better to express the terms' polarity based on a wide range of weight to increase the efficiency of the lexicon in storing and analyzing the sentiments.

4.2 Negation Lexicon

The negation lexicon is constructed for all Arabic negative words. For example, the terms " لا – لن – لم – ليس – مو – غير – مش - ما" are added to the negation lexicon and are used for both MSA and colloquial lexicons. During analysis process, the automatic lexicon analyzer will trace all the comment. If a negation word is detected, the term weight will be inversed. For example, the comment "موبايل **مش** كويس" means "not good mobile". If the term "كويس" or "good" is nearly positive of a weight value 1, it will be converted to nearly negative of a weight value -1.

4.3 Multi-Weighting Algorithm

The process of applying the multi-weighting polarity is effective in identifying and expressing all terms' polarity with high efficiency and flexibility. Different studies used the weighting process in Arabic sentiment analysis. The authors of [1] present an opinion mining and analysis for Arabic language but based on a fixed weight format. Fixed weight format identifies terms' polarity only as positive, negative, or neutral without defining the intensification degree of the term meaning especially in Arabic language. For example, the colloquial term "حلو" or "sweet" and the MSA term "جميل" or "beautiful". Both terms have positive polarity but the intensification degree of the term "جميل" or "beautiful" is higher than the term "حلو" or "sweet". The multi-weight polarity deals with this issue for increasing the flexibility and efficiency of both Arabic lexicon and sentiment analysis.

The authors of [2] presents a weighting methodology based on detecting some Arabic booster words like "كثير" which means "many or much" or "جدا" which means "very" for increasing the term weight but large number of Arabic sentiments can contain different terms that have variability in their polarity without having booster words.

For example; the comment " هذا اللاعب **موهوب** وله مستقبل **باهر**" or "This is a **talented** player and has a **bright** future". The terms "موهوب" which means "talented" and "باهر" which means "bright", have strong positive polarity without having booster words. As presented in [3], the weight of the term is based on dividing the number of polarity terms by the total number of the words in the text. So, for lengthy text, the degree of the term weight will decrease although the term polarity is strong. In addition, this algorithm is based on having only positive or negative terms in the comment but cannot deal with a comment that has positive and negative terms in the same comment.

One of the recent hybrid approaches used in sentiment analysis is presented in **[10]**. In this approach, data sets are analyzed based on their polarity to positive, negative, and neutral without explaining the polarity intensification of each term. In addition, the negation mechanism is not applied and the dictionaries are built manually. A number of thirteen dictionaries are used for storing data sets. This can increase the time complexity.

Another hybrid approach is presented in **[11]** for building lexicon-based and corpus-based sentiments. This paper depends on manual construction of lexicons. In addition, the classification of sentiment polarity is based on only positive, negative, and neutral terms.

In multi-weight polarity: the more the length of the sentiment text, the more the probability of having polarity terms. So, it was better to increase the weight of the sentiment polarity for increasing the flexibility of the lexicon and the efficiency of the sentiment analysis algorithm.

In this research, each positive or negative term has a specific weight whether it is in the same comment or in other comments. The average weight is calculated by dividing the sum of weights in the comment by the number of polarity terms in the same comment based on the following formula.

$$\forall \boldsymbol{t} \in \boldsymbol{s} \text{ such that } \boldsymbol{Avg_{weight}} = \frac{\sum w_i}{\sum t_i} \qquad \textbf{(1)}$$

Where $\boldsymbol{w_i}$ is the weight for each term $\boldsymbol{t_i}$.

The multi-weight polarity is based on 3-weight, 5-weight, and 7-weight. The 3-weight polarity is used for low intensification polarity; 5-weight is used for medium intensification polarity, while 7-weight is used for high intensification polarity. In each weight mechanism, five intervals are used for classifying the term polarity: positive (Pos), positive orientation (Pos_{ort}), neutral, negative orientation (Neg_{ort}), and negative (Neg).

Let

$\boldsymbol{Avg_{weight}}$: is the average weight of the terms in the sentence (S) calculated in formula (1).

S_p: is the general sentence polarity.

For a 3-weight polarity, the classification of terms is used as follows:

$$\boldsymbol{Min_{weight}} = -1$$
$$\boldsymbol{Max_{weight}} = 1$$
$$if\ (\boldsymbol{Avg_{weight}}) > 0\ and\ (\boldsymbol{Avg_{weight}}) \leq 0.5$$
$$then\ S_p \leftarrow Pos_{ort}$$
$$if\ (\boldsymbol{Avg_{weight}}) > 0.5\ and\ (\boldsymbol{Avg_{weight}}) \leq 1$$
$$then\ S_p \leftarrow Pos$$
$$if\ (\boldsymbol{Avg_{weight}}) = 0$$
$$then\ S_p \leftarrow Neutral$$
$$if\ (\boldsymbol{Avg_{weight}}) < 0\ and\ (\boldsymbol{Avg_{weight}}) \geq -0.5$$
$$then\ S_p \leftarrow Neg_{ort}$$
$$if\ (\boldsymbol{Avg_{weight}}) < -0.5\ and\ (\boldsymbol{Avg_{weight}}) \geq -1$$
$$then\ S_p \leftarrow Neg$$

For a 5-weight polarity, the classification of terms is used as follows:

$$\boldsymbol{Min_{weight}} = -2$$
$$\boldsymbol{Max_{weight}} = 2$$
$$if\ (\boldsymbol{Avg_{weight}}) > 0\ and\ (\boldsymbol{Avg_{weight}}) \leq 1$$
$$then\ S_p \leftarrow Pos_{ort}$$
$$if\ (\boldsymbol{Avg_{weight}}) > 1\ and\ (\boldsymbol{Avg_{weight}}) \leq 2$$
$$then\ S_p \leftarrow Pos$$
$$if\ (\boldsymbol{Avg_{weight}}) = 0$$
$$then\ S_p \leftarrow Neutral$$
$$if\ (\boldsymbol{Avg_{weight}}) < 0\ and\ (\boldsymbol{Avg_{weight}}) \geq -1$$
$$then\ S_p \leftarrow Neg_{ort}$$
$$if\ (\boldsymbol{Avg_{weight}}) < -1\ and\ (\boldsymbol{Avg_{weight}}) \geq -2$$
$$then\ S_p \leftarrow Neg$$

For a 7-weight polarity, the classification of terms is used as follows:

$$\boldsymbol{Min_{weight}} = -3$$
$$\boldsymbol{Max_{weight}} = 3$$
$$if\ (\boldsymbol{Avg_{weight}}) > 0\ and\ (\boldsymbol{Avg_{weight}}) \leq 1.5$$
$$then\ S_p \leftarrow Pos_{ort}$$
$$if\ (\boldsymbol{Avg_{weight}}) > 1.5\ and\ (\boldsymbol{Avg_{weight}}) \leq 3$$
$$then\ S_p \leftarrow Pos$$
$$if\ (\boldsymbol{Avg_{weight}}) = 0$$
$$then\ S_p \leftarrow Neutral$$
$$if\ (\boldsymbol{Avg_{weight}}) < 0\ and\ (\boldsymbol{Avg_{weight}}) \geq -1.5$$
$$then\ S_p \leftarrow Neg_{ort}$$
$$if\ (\boldsymbol{Avg_{weight}}) < -1.5\ and\ (\boldsymbol{Avg_{weight}}) \geq -3$$
$$then\ S_p \leftarrow Neg$$

5. NEGATION DETECTION ALGORITHM

The negation mechanism in Arabic sentiment analysis is based on storing different negation particles. If a negation particle is detected, the polarity of the term that appears after the negation particle will be converted to the inverse meaning. Different research papers mentioned the negation process during the sentiment analysis. The authors of [2, and 11] used a switch negation for reversing the meaning of the word. One of the recent researches in the negation of Arabic sentiment analysis is presented in [4]. In this research, the negation particles are separated into two groups. The first group contains four negation particles "ما – لم – لن - لا" which mean "not". These particles affect the verb that appears immediately after them. The second group contains the negation particle "ليس" which also means "not". This particle affects the following two nouns or affects the following verb. So, the negation terms are used to reverse the first word appears after them.

An improvement of the negation mechanism is presented in Algorithm 1. In this algorithm, negation method is not based on reversing the first verb appears after the negation particle because many sentiments contain negation particles that don not have polarity terms after them.

For example, the MSA comment "ليس عليك **ذنب**" or "You are not **guilty**" and the colloquial comment "أنت مش عليك أى **غلط**" or "You are never **wrong**". In these comments, the polarity term "ذنب" or "غلط" respectively have negative polarities and will not be reversed to positive polarities because they are not appeared immediately after the negation particle "ليس" or "مش" respectively. As a result, a false negative (FN) polarity will be found.

In addition, the Arabic exception method is used for handling a large number of Arabic sentiments that are analyzed with false positive or false negative polarities. The exception method is based on the exception negation particles ({ما - إلا} – {لا - إلا} – {مش - إلا}) which means "not - except" or "not - only". The terms {ما - لا – مش} are called negation particles. The term "إلا" is called exception particle.

For example, the colloquial comment "إحنا مش **خايفين** إلا من الأداء الإقتصادى" which means "We are not **afraid** only of economic performance". In this sentence, the term "خايفين" which means

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

"afraid" has a negative polarity. The traditional negation method will detect the negation particle "مش" which means "not". So, the term polarity will be reversed to positive one although the sentence has a negative polarity. This is a false positive polarity (FP). As result, the accuracy and the performance analysis will decrease.

Table 1: Exception Index

Remark	Sentence Polarity	n	...	k	...	Polarity Term (m)	...	i+1	i	...	2	1	Index
	Before Exception Negation												
FN	Negative					Positive							
FP	Positive					Negative							
	After Exception Negation												
				إلا					لا				
-	Positive					Positive							
-	Negative					Negative							

By applying the exception negation algorithm, the automatic lexicon will trace the sentence by identifying the index of each term. This is presented in Table 1. In Table 1, the exceptional negation algorithm is based on index structure. This index is generated for storing all Arabic sentiments by inserting the first term in the first index [i]. The next term is stored in [i+1] until the last term in the sentence that is stored in the index [n]. The overall sentiment analysis is presented in Algorithm 1.

Algorithm 1: Exceptional-Negation Detection

```
t ← Polar term
Nt1 ← Negation Term 1
Et2 ← Exceptional Term 2
Retrieve all Nt1 from Negation Lexicon
Retrieve all Et2 from Negation Lexicon
Trace all terms Nt1 and Et2 in MSA Lexicon
Trace all terms Nt1 and Et2 in Colloquial Lexicon
If count ( Nt1 ) > 0 and count ( Et2 ) > 0 then
   For each term Nt1 and Et2 ∈ MSA do
      Index [ t ] = m
      Index [ Nt1 ] = i
      Index [ Et2 ] = k
      If m > i and m < k then
         Item. Weight = Item. Weight * 1
      End If
   End For
   For each term Nt1 and Et2 ∈ Colloquial do
      Index [ t ] = m
      Index [ Nt1 ] = i
      Index [ Et2 ] = k
      If m > i and m < k then
         Item. Weight = Item. Weight * 1
      End If
   End for
Else If count ( Nt1 ) > 0 and count ( Et2 ) = 0 then
   For each term Nt1 ∈ MSA do
      Index [ t ] = m
      Index [ Nt1 ] = i
      If m > i then
         Item. Weight = Item. Weight * -1
      End If
   End For
   For each term Nt1 ∈ Colloquial do
      Index [ t ] = m
      Index [ Nt1 ] = i
      If m > i then
         Item. Weight = Item. Weight * -1
      End If
   End For
  End If
End If
```

As presented in Algorithm 1, in both MSA and colloquial lexicons, if exception particles are detected, the index of all terms in the sentence will be identified. The first negation particle will be stored in index [i] and the second exception particle will be stored in index [k]. If the index value of the polarity term [m] is greater than the index value of [i] and less than the index value [k], then the weight of the term will by multiplied by 1 in order to keep the item polarity without reversing. If exception particles are not detected, the automatic lexicon will trace the sentence and identifies the index value of the polarity term [m] and the negation particle [i]. If the index value of the polarity term [m] is greater than the index value of negation particle [i], then the weight of the term will by multiplied by -1 to reverse the polarity weight.

For example, the sentence " النظام ده مش هيكون **مفيد** إلا لو إتفعل من الصف الأول الثانوى" which means "This system will not be **useful** except if it is applied in the first secondary grade".

For a 5-weight polarity, the term "مفيد" which means "useful" has a positive polarity with weight value of +2. After the first negation particle "مش" or "not", the polarity weight of the term will be reversed to -2. If exceptional particle "إلا" or "except" is detected, the polarity weight of the term will be reversed back to +2. This can eliminate any false positive (FP) or false negative (FN) polarities. In addition, it is not required that the polarity verb or noun must appear immediately after the negation particle. This is considered an improvement of the negation methods presented in [2] and [4].

6. EXPERIMENTAL RESULTS

The automatic lexicon with the exception negation algorithm is developed using Microsoft Visual Studio.net 2015 with Microsoft SQL Server 2010 database. The experimental results were conducted on Intel ® Core i5 @ CPU 1.8 GHz machine with 4GB RAM. The operating system was Microsoft Windows 10. As presented in Table 2, a total of 3,476 posts are collected from Facebook, Twitter, and movie reviews [22] distributed in four categories: economy, sport, history, and politics. These datasets are stored in the automatic lexicons and tested for multi-domain data with different weight values.

Table 2: Data set Distribution

	Economy	Sport	History	Politics
No of Posts	852	884	988	752

After applying the multi-weight polarity and the exceptional negation algorithm in the automatic lexicon, all sentiments are analyzed based on their polarity whether they are: positive, nearly positive, neutral, nearly negative, and negative. The resulted sentiment polarities are tested using Rapidminer 7.3.001 machine learning tool to test the overall performance of the results based on accuracy, precision, recall, and F-measure.

The experimental results are conducted using the three classifiers: Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and Naïve Bayes (NB). In the cross-validation process of the three classifiers, different splitting ratios are examined for splitting the data into training set and testing set. The best splitting validation ratio that achieves the highest accuracy is set to 0.7. For KNN algorithm the value of k =3 is set during the analysis process. The accuracy, precision, and recall are recorded based on the three polarity weight W_3, W_5, and W_7.

Table 3: Accuracy using Polarity Weight W_3

	Economy	**Sport**	**History**	**Politics**
KNN	0.952	0.949	0.959	0.964
SVM	0.983	0.98	0.966	0.927
NB	0.947	0.98	0.973	0.964

Table 4: Precision using Polarity Weight W_3

	Economy	**Sport**	**History**	**Politics**
KNN	0.961	0.962	0.884	0.967
SVM	0.979	0.982	0.96	0.938
NB	0.95	0.981	0.89	0.93

Table 5: Recall using Polarity Weight W_3

	Economy	**Sport**	**History**	**Politics**
KNN	0.964	0.953	0.9074	0.985
SVM	0.985	0.9821	0.967	0.925
NB	0.95	0.981	0.918	0.952

As presented in Table 3, SVM achieves high accuracy rate in W_3 with 98.3% in economy dataset. SVM and NB achieve an equal accuracy of 98% in sport while NB achieves 97.3% in history dataset. KNN and NB achieve the same accuracy in politics dataset with 96.4%. In Table 4, the precision of W_3 is measured where SVM achieves 97.9%, 98.2%, and 96% in economy, sport, and history datasets respectively. KNN achieves better precision with 96.7% in politics dataset. In Table 5, SVM achieves better recall in economy, sport, and history datasets with 98.5%, 98.21%, and 96.7% respectively while KNN achieves better recall in politics dataset with 98.5%. The precision and recall of NB in W_3 have low rates when it is compared with SVM and KNN.

Table 6: Accuracy using Polarity Weight W_5

	Economy	**Sport**	**History**	**Politics**
KNN	0.9688	0.9712	0.985	0.9643
SVM	0.9483	0.9704	0.9681	0.9545
NB	0.9844	0.9723	0.958	0.9821

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 7: Precision using Polarity Weight W_5

	Economy	Sport	History	Politics
KNN	0.971	0.968	0.9722	0.9467
SVM	0.971	0.96	0.9844	0.9783
NB	0.962	0.968	0.9889	0.9333

Table 8: Recall using Polarity Weight W_5

	Economy	Sport	History	Politics
KNN	0.984	0.98	0.975	0.9646
SVM	0.9516	0.98	0.9619	0.9616
NB	0.992	0.98	0.9544	0.98

As presented in Table 6, NB achieves better accuracy in W_5 with 98.44%, 97.23%, and 98.21% in economy, sport, and politics datasets respectively, while KNN achieves better accuracy with 98.5% in history dataset. In Table 7, the precision of W_5 is measured where NB achieves high precision in both economy and politics datasets with 96% and 97.5% respectively. In sport dataset, SVM, KNN, and NB have the same precision with 95.1%. In history dataset, KNN achieves better precision with 95.7%. In Table 8, the same performance of algorithms in recall is similar to precision where NB achieves high recall in both economy and politics datasets with 99.2% and 98% respectively. In sport dataset, SVM, KNN, and NB have the same recall with 98%. In history dataset, KNN achieves better recall with 97.5%.

Table 9: Accuracy using Polarity Weight W_7

	Economy	Sport	History	Politics
KNN	0.975	0.9697	0.973	0.9286
SVM	0.975	0.9623	0.9825	0.9778
NB	0.9594	0.9697	0.9865	0.9107

Table 10: Precision using Polarity Weight W_7

	Economy	Sport	History	Politics
KNN	0.975	0.9735	0.9789	0.9346
SVM	0.975	0.9667	0.9808	0.9783
NB	0.9676	0.9735	0.9889	0.9231

Table 11: Recall using Polarity Weight W_7

	Economy	Sport	History	Politics
KNN	0.9505	0.951	0.957	0.9583
SVM	0.95	0.951	0.9343	0.95
NB	0.96	0.951	0.9291	0.975

As presented in Table 9, SVM and KNN achieve better accuracy in W_7 with 97.5% in economy dataset, while KNN and NB achieve better accuracy with 96.97% in sport dataset. NB achieves high accuracy with 98.65% in history dataset while SVM achieves high accuracy in politics dataset with 97.78%. In Table 10, the precision of W_7 is measured where SVM and KNN achieve high precision in economy set with 97.5%. In sport dataset, KNN and NB achieve better precision with 97.35%. In history dataset, NB achieves better precision with 98.89% while SVM achieves high precision in politics dataset with 97.83%. In Table 11, the same performance of algorithms in recall is similar to precision and accuracy where SVM and KNN achieve high recall in economy dataset with 97.1%. In sport dataset, KNN and NB achieve better recall with 96.8%. In history dataset, NB achieves better recall with 98.89% while SVM achieves high precision in politics dataset with 97.83%.

The average accuracy, precision, recall, and F-measure using the multi-weight algorithms W_3, W_5, and W_7 are presented in Figure 1, 2, 3 and 4. The experimental results are measured using SVM, KNN, and NB and are compared with the experimental results presented in most recent research papers [5, 6, 7, and 20] to determine the overall performance of the system.

The authors of [5] propose a framework for combining sentiment analysis with subjective analysis to determine whether people are interested in defined subject or not. The authors of [6] evaluate the Senti-Strength about their adaptability for Arabic sentiments. The authors of [7] perform an analysis of sentiments by implementing a framework for determining the polarity of terms. Finally, the authors of [20] build a lexicon based on two different Arabic corpora: book review corpus and OCA corpus. The book review corpus was developed by crawling several book reviews websites and manually determines their sentiment polarity. The OCA corpus [22] is a movie review corpus consisting of 250 positive and 250 negative movie reviews in Arabic.

As presented in [5], the average accuracy after stemming process using SVM and NB achieve 87.7996% and 85.1872% respectively while the accuracy presented in [6] after evaluating Senti-Strength achieves only 78.1%. As explained in [20], the accuracy of book review corpus using NB and SVM achieves 94.88% and 76.69% respectively while the accuracy of OCA corpus using NB and SVM achieves 97.81% and 89.29% respectively.

After applying the multi-weight algorithms with exceptional negation detection algorithm presented in this paper, better accuracy is achieved as explained in Figure 1. In Figure 1, NB and SVM on W_3 achieve average accuracy of 96.618% and 96.37% respectively better than KNN that achieves 95.648%. On W_5, NB and KNN achieve a relative accuracy of 97.42% and 97.233% respectively better than SVM that achieves 96.033%. On W_7, SVM achieves high accuracy of 97.44% which is better than KNN and NB that achieve 96.158% and 95.658% respectively.

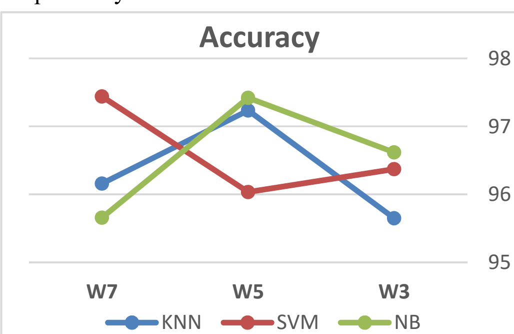

Figure 1: Average System Accuracy

As presented in [5], the average precision using SVM and NB achieve 79% and 88% respectively while the precision presented in [6] achieves 92.9%. As presented in [7], the precision achieves 91.2% using SVM, 87.6% using NB, and 76% using KNN using Weka tool.

After applying the multi-weight algorithms with exceptional negation detection algorithm presented in this paper, better precision is achieved as explained in Figure 2. In Figure 2, SVM on W_3 achieves high precision of 96.643% greater than KNN and NB that achieve 94.368% and 93.768% respectively. On W_5, KNN and NB achieve a relative precision of 95.42% and 95.378% respectively better than SVM that achieves 94.633%. On W_7, SVM achieves high accuracy of 97.44% which is better than KNN and NB that achieve 96.55% and 96.328% respectively.

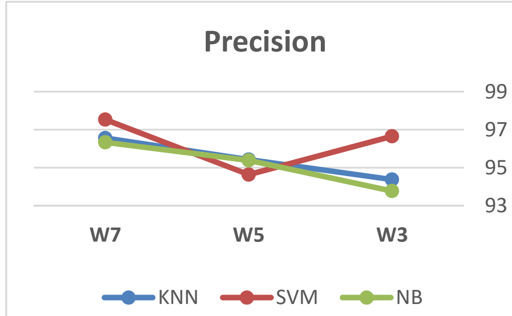

Figure 2: Average System Precision

The average recall presented in [5] achieves 96.4% and 85.2% in SVM and NB respectively while the experimental results presented in [6] achieves low recall of 81.5%. As presented in [7], the recall achieves 90.2% using SVM, 87.3% using NB, and 70.5% using KNN using Weka tool.

The multi-weight algorithms with exceptional negation detection algorithm presented in this paper achieve better recall as explained in Figure 3. In Figure 3, SVM on W_3 achieves high recall of 96.478% greater than KNN and NB that achieve 95.22% and 95.024% respectively. On W_5, NB and KNN achieve very relative recall of 97.66% and 97.59% respectively better than SVM that achieves 96.378%. On W_7, SVM achieves high recall of 97.343% which is better than KNN and NB that achieve 96.448% and 96.303% respectively.

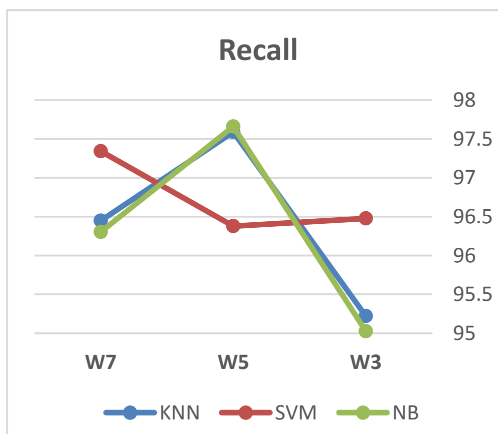

Figure 3: Average System Recall

The F-measure results presented in [5] achieves low rates of 86.8% using SVM and 85.3% using NB. As presented in [6], the largest F-measure of all data sets is 86.8%.

The F-measure is calculated for the multi-weight algorithms and achieves high rates compared with experimental results presented in [5] and [6]. As presented in Figure 4, SVM on W_3 achieves high F-measure of 96.56% better than KNN and NB that achieve 94.792% and 94.392% respectively. On W_5, NB and KNN achieve F-measure of 96.51% and 96.493% respectively better than SVM that achieves 95.498%. On W_7, SVM achieves F-measure of 97.431% which is better than KNN and NB that achieve 96.499% and 96.315% respectively.

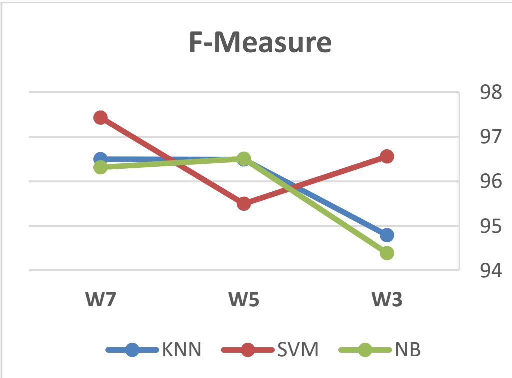

Figure 4: Average System F-Measure

From the experiments presented in this paper, the results show an improvement of SVM over KNN and NB in precision, recall and F-measure on both W_3 and W_7. In the accuracy, SVM achieves better results on only W_7 while on W_3 NB achieves better accuracy than SVM although SVM accuracy is relative to NB accuracy. On W_5, both NB and KNN achieve relative results better than SVM in accuracy, precision, recall, and F-measure.

7. CONCLUSION AND FUTURE WORKS

The wide increase of social networks and reviews make the use of sentiment analysis more important. Analysis of sentiment can be very beneficial for businesses, governments, and individuals. This paper presents an automatic lexicon for modern standard Arabic (MSA) and colloquial sentences that can deal with classical, slang, and mixed sentiments. A negation detection algorithm is enhanced in this paper for reducing false positive and false negative rates that result from the wrong polarity identification. The experimental results were conducted using SVM, KNN, and NB that explain an improvement over recent sentiment analysis researches. The future work of this paper is based on two directions. The first one is to increase the size of data sets to ascertain the effectiveness of presented Algorithm. The second one is to apply the n-gram computational linguistics in the analysis process to distribute the polarity term analysis.

REFRENCES

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

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

[3] S. R. El-Beltagy, A. Ali, “Open Issues in the Sentiment Analysis of Arabic Social Media: A Case Study,” *IEEE International Conference on Innovations and Information Technology (IIT)*, 2013, pp.1-6

[4] R. Duwairi, and M. Alshboul, "Negation-aware Framework for Sentiment Analysis in Arabic Reviews," *IEEE International Conference on Future Internet of Things and Cloud*, 2015, pp. 731-735

[5] N. Bin Haithlian, and A. Hafez, "Sentiment – Subjective Analysis Framework for Arabic Social Media Posts," *IEEE Saudi International Conference on Information Technology*, 2016, pp. 1-6

[6] A. Rabab'ah, M. Al-Ayyoub, Y, Jararweh, and M. Al-Kabi, "Evaluating SentiStrength for Arabic Sentiment Analysis," *IEEE International Conference on Computer and Information Technology*, 2016, pp.1-6

[7] M. Sghaier, and M. Zrigui, "Sentiment Analysis for Arabic e-Commerce Websites," *IEEE International Conference on Engineering and MIS*, 2016, pp.1-7

[8] B. Hawashin, A. Mansour, and S. Aljawarneh, "An Efficient Feature Selection Method for Arabic Text Classification," *International Journal of Computer Application*, vol. 83, no. 17, 2013, pp. 1- 6

[9] A. Mostafa, "An Evaluation of Sentiment Analysis and Classification Algorithms for Arabic Textual Data," *International Journal of Computer Application*, vol.158, no.3, 2017, pp.29-36

[10] R. Khasawneh, H. Wahsheh, I. Alsmadi, and M. Al-Kabi, "Arabic Sentiment Polarity Identification using a Hybrid Approach," *IEEE International Conference on Information and Communication Systems*, 2015, pp.148-153

[11] N. Abdulla, N. Ahmed, M. Shehab, and M. Al-Ayyoub, "Arabic Sentiment Analysis:

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

lexicon-based and Corpus-based," *IEEE International Conference on Applied Electrical Engineering and Computing Technologies*, 2013, pp. 1-6

[12] W. Medhat, A. Hassan, and H. Korashy, "Sentiment Analysis Algorithms and Applications: A Survey" *Ain Shams Engineering Journal, ELSEVIER*, vol.5, 2014, pp.1093-1113

[13] J. Akaichi, "Sentiment Classification at the Time of the Tunisian Uprising," *IEEE International European Conference on Network Intelligence*, 2014, pp. 38-45

[14] Q. Al-Radaideh, and L. Twaiq, "Rough Set Theory for Arabic Sentiment Classification," *IEEE International Conference on Future Internet of Things and Cloud*, 2014, pp.559-564

[15] M. Faqeeh, N. Abdulla, Y. Al-Ayyoub, and M. Quwaider, "Cross-lingual Short-text Document Classification for Facebook Comments," *IEEE International Conference on Future Internet of Things and Cloud*, 2014, pp.573-578

[16] R. Duwairi, "Sentiment Analysis for Dialectical Arabic," *IEEE International Conference on Information and Communication Systems*, 2015, pp.166-170

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

[18] H. Ibrahim, S. Abdou, and M. Gheih, "MIKA: A Tagged corpus for Modern Standard Arabic and Colloquial Sentiment Analysis," *IEEE International Conference on Recent Trends in Information Systems*, 2015, pp.353-358

[19] D. Mohey El-Din, "A Survey on Sentiment Analysis Challenges," *Journal of King Saud University – Engineering Sciences, ELSEVIER*, 2016, pp.1-9

[20] F. Mayhoub, M. Siddiqui, and M. Dahab, "Building an Arabic Sentiment Lexicon using Semi-Supervised Learning," *Journal of King Saud University – Computers and Information Sciences, ELSEVIER*, vol.26, 2014, pp.417-424

[21] J. Guerrero, J. Olivas, F. Romero, and E. Viedma, "Sentiment Analysis: A Review and Comparative Analysis of Web Services," *Journal of Information Sciences, ELSEVIER*, vol. 311, 2015, pp.18-38

[22] M. Saleh, M. Valdivia, L. Lopez, and J. Ortega, "OCA: Opinion Corpus for Arabic," *Journal of the Association for Information Science and Technology, WILEY*, vol. 62, 2011, pp.2045-2054