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

PERFORMANCE EVALUATION OF AN ADOPTED SENTIMENT ANALYSIS MODEL FOR ARABIC COMMENTS FROM THE FACEBOOK

IBRAHIM ROUBY*, MOHAMMED BADAWY, MOHAMED NOUR*, NADIA HEGAZI***

Electronics Research Institute, Cairo, Egypt*
Faculty of Electronic Engineering, Menoufia University, Menouf, 32952, Egypt**
E-mail: [1]ibrahimelgazwy@gmail.com

ABSTRACT

Nowadays, the resources of social media are important for sharing data, news, and opinions. Users of social media can write their tweets, posts, and comments to express their feedback about some services and products. Sentiment analysis is one of the approaches for analyzing users` opinions to extract useful information. This research work analyzes and investigates a sentiment analysis model. The model contains four phases mainly: document/dataset collection, preprocessing operations, scoring and sentiment classification, and evaluation. The dataset collection is concerned with collecting Arabic documents or comments from social media like Facebook. The preprocessing operations involve tokenization, rejection of stopwords, normalization, and stemming. Scoring and sentiment classification are concerned with many important themes mainly: checking negation, handling intensifiers, identifying emotions and sentiment classification. The evaluation phase evaluates the performance of the sentiment analysis model. Moreover, the sentiment analysis model is supported by a set of Arabic lexical resources such as list of Arabic stopwords, list of positive and negative emotions, list of positive and negative modifiers, list of affixes of the light stemmer, and others. The sentiment analysis model helps classifying the users` comments to either positive or negative or neutral sentiments (Sentiment Polarity). The adopted sentiment analysis model is presented to identify sentiments in the Modern Standard Arabic (MSA). The model also can investigate and identify sentiments in informal Arabic (colloquial) where most of social media users are using. Some measurable criteria such as precision, recall, accuracy, and error-rate are adopted to evaluate the performance of the sentiment analysis model. Several experiments are done adopting three important themes of Arabic words mainly: negations, emotions, and intensifiers. The model behavior is changed and affected by using such themes either individually or combined. The model performance is also affected by using the type of Arabic sentence and Arabic language style. Finally, the sentiment analysis model behaves well and presents good accuracy values. The accuracy values of the predicted positive comments are 98.2%, 91.8%, and 85.8% while the values are 93.2%, 92.6%, and 70.1% for the negative comments respectively for MSA, Mixed Arabic, and informal Arabic styles.

Keywords: *Sentiment Analysis, Sentiment Polarity, Social Media, Arabic Text, Sentiment Classification.*

1. INTRODUCTION AND RELATED WORK

Sentiment analysis and/or opinion mining is one of the important research topics of computational linguistics. The sources of social media are important for sharing news, ideas, information and opinions. The companies, organization, and governments need to analyze the users` comments to obtain some knowledge for improving their decision making process. Sentiment analysis is important as it determines if the users` comments have positive or negative polarities. The majority of sentiment analysis research works are focused on the English language, while little are conducting those posts and comments written in Arabic. Examples of the efforts that handled the sentiment analysis techniques for Arabic datasets are briefly mentioned as follows:

[3] focused on sentiment analysis for Arabic tweets. The authors applied some supervised machine learning approaches to classify those tweets as positive or negative polarities. They applied some preprocessing techniques to improve the sentiment performance. The preprocessing techniques achieved better results.

[2] discussed the challenges and obstacles of the sentiment analysis for informal Arabic and social media. The authors mentioned that most of social media networks such as Tweeter and YouTube websites are using informal Arabic. The main differences between Tweeter and Youtube dialect were presented. Some methods and preprocessing operations were also adopted.

[4] mentioned that sentiment can be analyzed and classified either by machine learning techniques or by lexicon-based techniques. Tweeter is considered one of the most popular social networking where millions of users share their suggestions and opinions about several fields. Examples of such fields are: politics, products, personalities, and others. Sentiment analysis is used for interpreting the resources of the sentiment change in public attitude, mining, and summarizing products reviews to solve the polarity shift problem. The authors used different classification models like Naïve Bayes classifier, support vector machine, decision tree, and others.

[7] mentioned that the online contents in Arabic are limited and the related accuracy of the existing methods is lower. Arabic sentiment analysis is based on two main issues: the Arabic specific challenges and general linguistic issues. The Arabic specific challenges are caused by Arabic morphological complexity, limited resources, and dialects. The general linguistics issues include polarity fuzziness, spam, review quality, and others.

[19] mentioned that there are several approaches for detecting sentiment in text. Lexical resources; such as a dictionary of opinion terms; are important. SentiWordNet is a resource containing opinion information on terms extracted from the WordNet database. Using semi-supervised learning methods enable users to classify products` reviews as positive or negative using SentiWordNet lexicon.

[12] presented a lexicon-based sentiment analysis algorithm and focused on real-time tweeter content analysis. This is necessary to estimate the intensity of the sentiment rather than positive/negative label. The authors mentioned that sentiment analysis has been applied on different, non-necessary domains for monitoring and forecasting public opinions. The authors presented and explained the relation between negative sentiment of Tweeter and posts related to English Defense League and the level of disorder during the organization's related events.

[13] tested the effects of normalization, stemming, and stop-words removal on the performance of Arabic sentiment analysis system. The system used Arabic tweets from Tweeter. The sentiment of the crawled tweets was analyzed to interpret the attitude of the public. The authors mentioned that determining the writer's attitude regarding some topics is the main task of sentiment analysis.

[24] presented an approach to sentiment polarity classification in twitter posts. The approach is based on extracting a vector of weighted nodes from the graph of WordNet. The weights are used in SentiWordNet to compute a final estimation of the polarity. The approach is considered a non-supervised solution that is domain independent. The performance evaluation of their work shows that the approach is promising.

[25] mentioned that the advances in Web 2.0 had changed the ways people communicate, collaborate, and express their opinions and sentiments. The authors developed SenticNet2, a publicly available semantic and affective resource for sentiment analysis. SenticNet2 was built by means of sentic computing which is a paradigm that exploits both artificial intelligence and semantic web techniques to better recognize, interpret, and process natural language opinions. SenticNet2 is considered one of the most comprehensive semantic resources for the development of affect-sensitive applications in several areas such as data mining, multimodal HCI, social media marketing, and others.

[26] introduced some aspects of web data analysis. They mentioned that the process of sentiment analysis involves including the identification of the meaning of words within the text through natural language processing rules. The authors also presented a framework for automatic identification of the presence of opinion in textural data. The framework includes a description of rules of negation identification and calculation. The negation rules are designed to improve sentiment text analysis.

[27] introduced an approach that automatically classifies the sentiment of tweets by using classifier ensembles and lexicons. Classification of tweets is done concerning a query term. This work is useful for customers who can use sentiment analysis to search for products, for companies that aim at monitoring the public sentiments of their brands and for other applications. The authors' experiments on some public tweet sentiment datasets show that

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

classifier ensembles formed by Naïve Bayes, SVM, and logistic regression improved the classification accuracy.

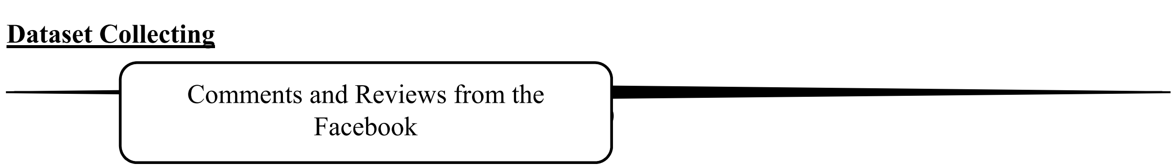
[28] mentioned that the reviews and blogs obtained from social networks are very important resources for analysis and improving decision making. The reviews are unstructured by nature and need some sort of processing like classification and clustering to provide useful information for future uses. Supervised machine learning methods help to classify such reviews. The authors adopted four different machine learning algorithms for classifying the human sentiments. The algorithms are: Naïve Bayes, SVM, Maximum Entropy, and Stochastic Gradient Descent. They considered some measurable criteria for evaluating the performance of the adopted algorithms.

[29] mentioned that determining the sequence of words affected by negation is one of the important tasks of sentiment analysis. The authors integrated the problem of identifying the scope of negation while determining the polarity of sentence. They proposed a negation handling approach based on linguistic features which determine the effect of different types of negation. The proposed approach improved the accuracy of negation identification and overall sentiment analysis.

The organization of this work will be as follows: Section 2 presents the architecture of the adopted sentiment analysis model. This includes several key points mainly: the dataset collection, preprocessing operations, scoring and sentiment calculation, and performance evaluation. The sentiment calculation involves negation handling, intensifier checking, and emotion identification. Section 3 presents a proposed approach for negation checking and handling. Section 4 presents the implementation work while Section 5 discusses the obtained results. Finally, the concluding remarks are presented in Section 6.

2. ARCHITECTURE OF THE ADOPTED SENTIMENT ANALYSIS MODEL

This work presents the architecture of the adopted sentiment analysis model. The architecture contains four main phases namely: dataset collection, preprocessing operations, scoring and sentiment classification, and performance evaluation. The architecture is also supported by a set of important lists to facilitate the calculation of sentiment scoring. The lists are: Arabic Stopwords list, list of Arabic affixations of the light stemmer, list of emotion symbols, list of Arabic negation words, and a part of Arabic lexicon. Figure 1 briefly shows the main themes of the sentiment analysis model, where Figure 1.1 shows the adopted Arabic lexical resources.



Journal of Theoretical and Applied Information Technology
15th November 2018. Vol.96. No 21

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

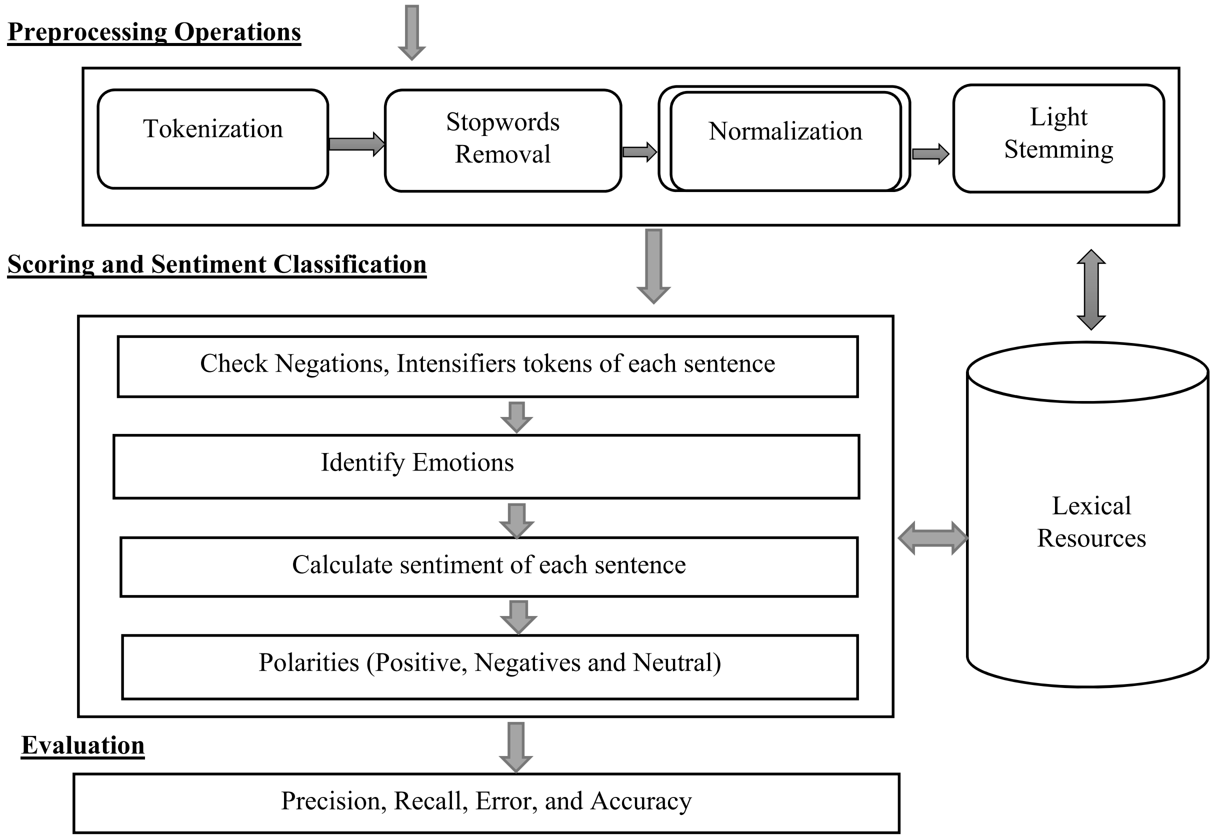


Figure. 1. Architecture of the Adopted Sentiment Analysis Model

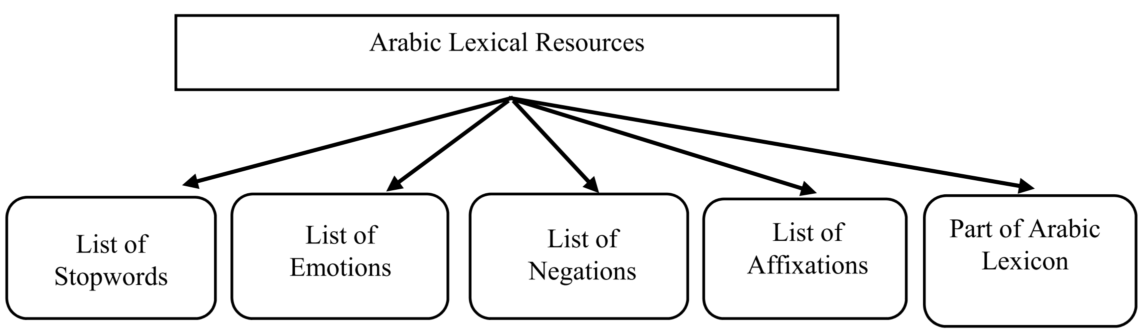


Figure. 1.1. The Arabic Lexical Resources Used in the Adopted Model

2.1. Dataset Collection

To operate and test a dataset for sentiment analysis, it is normal to adopt three steps mainly: perform web crawling, filter the reviews, and extract some attributes from each reviewer. The collection of data is important to compile the dataset from users` reviews. The dataset is considered as input to the preprocessing phase. In our case, we have three datasets; each contains a number of comments. The datasets were collected from the Facebook API that offers a secure HTTP-based API. It allows developers to query public posts of specific users or organizations via authenticated HTTP calls. The colloquial dataset collection represents comments about different events and news from Youm7 page on Facebook. The number of colloquial comments is 1402. The Modern Standard Arabic (MSA) represents a sample of Large Scale Arabic Book Reviews (LSABR). The number of MSA comments is 341.

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

2.2. The Preprocessing Operations

This phase is focused on text preprocessing which is important for sentiment analysis. This phase presents several operations mainly: tokenization, stopwords' rejection, normalization, and light stemming. The function of each operation is briefly mentioned as follows:-

2.2.1 Tokenization of Arabic Text

Tokenization is an important step and is performed for each comment in the adopted dataset. Tokenization divides and/or splits any comment into multiple tokens or words based on whitespace character or any delimiter. The tokens may be number, article, preposition, modifier, punctuation, or any word of different types (nouns, verbs, adjectives, adverb, …etc.).

Table 1: Examples of Some Arabic Stopwords

عن	أمام	إلى	تلك	التى	هذا
كانوا	خلف	على	ذلكما	اللذان	هذه
كن	كان	فى	تلكما	اللتان	هذان
هو	كانت	حتى	ذلكم	اللآتى	هاتان
هى	كانا	فوق	تلكم	اللآئى	هؤلاء
هم	كانتا	تحت	من	ذلك	الذى
…	…	…	…	…	…

2.2.2 Rejection of Stopwords

Stopwords are those words that are not important in calculating a term weighting. The stopwords are rejected or removed as they do not have any effect on identifying the sentiment classification. The stopwords are sometimes called noise words and they may be prepositions, pronouns, modifiers, specifying tools, and other tools. The stopwords mentioned in Table 1 are considered a part of stopwords exist in the Arabic Language (MSA and Colloquial). The original list of Arabic stopwords contains about 607 stopwords.

2.2.3 Normalization

Normalization is important for a good accurate sentiment analysis model. Normalization is concerned with removing a set of letters. This includes removing punctuations, weak vowels, removing elongations (or repeated characters), and removing non-letters like numbers. Normalization also is concerned with replacing some letters by others. Letters like (alef madda آ), (alef Hamza أ), and (alef kasra إ) can be replaced with (alef ا). The same thing also can take place by replacing final (taa marbuta ة) by (haa marbuta ه) and so on. [14], [13], [2].

2.2.4 The Arabic Light Stemming

Light stemming is simple and proves to be effective in a lot of computational linguistic applications. A light stemmer is adopted to suit the task of sentiment analysis. The light stemmer removes both the Arabic prefixes and the Arabic Affixes.[14], [13], [2]. Table 2 shows examples of some Arabic prefixes, suffixes, and affixes.

Table 2: Examples of Some Arabic Prefixes, Suffixes, and Affixes

Prefix	Prefix	Suffix	Suffix	Affix	Affix
الدنيا	بلاكراه	مفلسون	اكلنا	المفسدون	اشفقنا
يعمل	ناكل	صالحين	شربوا	ليقولون	يتكلمون
فاليعلم	يستسلم	كتابك	معطفك	فادخلوا	يستغفرون
كالصخر	استغفر	لعبتكم	سبيلكم	بايديكم	نشكركم
…	…	…	…	…	…

2.3. Scoring and Sentiment Calculation

This step involves several important themes for the adopted sentiment analysis approach. This includes negation checking and handling, intensifier checking, emotion identification, and calculation of sentiment polarities. Such themes are analyzed and discussed as in the following subsections.

2.3.1 An Adopted Method for Negation Handling

Negation plays an important role in computing the sentiment scoring. The common methods for negation management can take place by using a lexicon. That lexicon contains a set of words; each has a corresponding sentiment value. The Arabic sentiment like those ones written in other natural languages (e.g. English, French, Italian …) may contain some negation terms. The negation terms can change the sentiment value of the sentence. For example, the sentence "الرجل كريم", "الرجل غير كريم" have different polarities. The first sentence has a positive sentiment while the second one reverses the polarity of the opinion word "كريم" from positive to negative. Adopting the lexicon approach needs to create a list containing the negation terms. Each word in a sentence is checked against that negation list Neg. i.e

$$\text{Neg} = \{\text{set of negation words}\}$$

If a word is found in the negation list, the polarity of the neighboring opinion word is computed by multiplying the score of opinion word by -1 as in Equation (1).

$Pol_{score\text{-}neg}(w_x) = \{ pol_{score}(w_x) * (-1), \text{ if } (w_x\text{-}1) \in Neg \}$ (1)

Where wx denotes the neighboring opinion word and wx-1 is the preceding word of an opinion word which belongs to the negation list Neg. Some researches recommended the idea of the reversing the polarity of the lexicon word for handling the negation as mentioned above. Other researches on the other hand adopted a different approach to calculate the sentiment value of the negation word by using a negation function. [12] adopted a negation function as shown in Equation (2).

$$F_N(S) = \begin{cases} \text{Max}\left\{\frac{s+100}{2}, 10\right\} & \text{If } S < 0 \\ \text{Min}\left\{\frac{s-100}{2}, -10\right\} & \text{If } S > 0 \end{cases}$$

The sentence becomes 10. [12] attempted to design a sentiment combining function that provides the absolute sentiment of the comment as normalized value from the range of -100 to 100. The normalized formula combines the average sentiment of the sentence and the number of words to calculate the average. The difference between the overall positive and overall negative sentiments depends on the number of positive and negative words in the comment. The authors determined the following normalization formula as shown in equation (3).

$$F_P = min\left\{\frac{A_P}{2 - log(3.5 \times W_P + I_P)}, 100\right\}$$
$$F_N = max\left\{\frac{A_N}{2 - log(3.5 \times W_N + I_N)}, -100\right\}$$

Where I_p and I_N are the member of intinsifiers that refer respectively to positive and negative words in a sentence. For more details, the reader can refer to [12], [8].

2.3.2 Checking of Intensifiers

A comment may include a word(s) called intensifier(s). Intensifiers exit in English, Arabic and other natural languages. Intensifiers can increase or decrease the intensity of a sentiment. In English, for example, words like ‘very’, ‘quite’, and ‘most’ can change the sentiment of the neighboring non-neutral terms. In Arabic, words like ‘جدا’, ‘قليلا’, ‘بشده’can also change the sentiment of the neighboring non-neutral Arabic words. Intensifiers can be of two types: amplifiers and downtoners where they can increase or decrease the intensity of sentiment respectively. Some researchers assign different values of sentiment according to the degree of amplification or downtoners. In our case and for simplicity, amplifiers and downtoners can increase and decrease the sentiment value by 50% and -50% respectively regardless the degree of weakness or strong of any intensifier item. A set of Arabic intensifier words were selected focusing on the most frequently applied ones. Intensifier items include but not limited to: بشده , غالبا , كثيرا , قليلا , جدا [12], [8].

2.3.3 Identification of Emotions

Emotions play an important role in calculating the sentiment score. Emotions are used in social media chatting. Emotion is a symbol representing the states of mind mood or feeling. Emotions are considered effective messages domain independent and language independent as well. As emotions are used on a large scale when writing comments, it is important to detect emotions in the user`s comments. Some researchers collected different symbols of emotions and put them in a dictionary or positive and negative lists. This will facilitate detecting the emotion symbols in a dataset. Examples of some collected emotion symbols are illustrated in Table 3. It is easy to say that emotions may be positive or negative. Any emotion can be labeled as positive if it is found in the positive list Ep. An Emotion is labeled as negative if it is found in the negative list En. The positive and negative lists can be written as follows:

(3) E_p = { list of positive emotions }. (5)

E_n = { list of negative emotions }.

The sentiment score and/or polarity is computed as shown in (6):

$$Pol_{score}(e) = \begin{cases} -1 & \text{if } e \in E_n \\ 1 & \text{if } e \in E_p \\ 0 & \text{if } e \notin E_p \cap e \notin E_n \\ 0 & \text{if } e \in E_p \cap e \in E_n \end{cases} \quad (6)$$

Table 3: Examples of Some Positive and Negatives Emotions

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

Emotion	Icon	Code	Sentiment Class
Smile		:-)	Positive
Cool		8-)	Positive
Happy		:)	Positive
Relaxed		:o	Positive
Sad		:(	Negative
Angry		:-||	Negative
Down		:o)	Negative

3. THE PROPOSED APPROACH FOR HANDLING ARABIC NEGATION WORDS

As mentioned before, negation is an important theme as it changes the meaning if it is used within a sentence. Negation words are those words which affect the sentiment orientation of other words in sentence. Negation can invert the polarity of opinionated words. Negation words are not easy as they are difficult to identify which part of a sentence is changed. Negation may occur more than one time in a sentence and a sentiment can cancel each other. In some natural languages (like English for example), there are three classes of negations mainly: syntactic negations, diminished negations, and morphological negations. The syntactic negations include those words like no, not, wasn't, didn't, haven't, none, cannot,…etc. Diminished negations are those words that change the intensity of words like less, little, hardly, etc. The morphological negations are the affixation such as de-, dis-, in-, mis-, im-, less-, … etc. In this work only syntactic negations are handled for Arabic comments. There are no morphological negations in Arabic sentences. Examples of Arabic negation words are ما, لا, لم, لن , which are formal negations. The informal Arabic contains some informal negation words. Such negation words differ from a society to another. E.g the word "موش" is a common negation word in the Egyptian society. A negation word like "مو" is a common negation word in the gulf societies. To find out the scope of the negation, all the words in any sentence should be identified. Negations can be analyzed using several aspects; examples of such aspects include but not limited to the part of speech and dependency tree. Combining both the part of speech and dependency tree are important in polarity calculation. The polarity of a sentence is also based on the meaning of its words. The dependency tree indicates how negation is interacting with other words in the sentence. To calculate any sentence polarity it is necessary to consider all the sentence constructs. The constructs may be noun, verb, objective, pronoun, etc. [26]. Polarity calculator gets the values of all noun phrases and also the verbs and adverbs of the verb phrase. The values of such items are reversed in case of negation. Polarity calculation can be done by considering the following examples; as shown in Figure 2; supported by their dependency trees respectively.

Example 1: Verbal Sentence
"نجح الطالب فى الاختبار "
VS: VP NP
VP: -V
NP: -N PP
PP: - Prep N

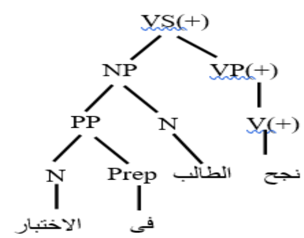


Fig. 2a. The Dependency Tree of a Verbal Sentence

Example 2: Negated Verb
"لا يكذب المؤمن فى حديثه"
VS: VS NP
VP: -Neg V
NP: -N PP
PP: -Prep N

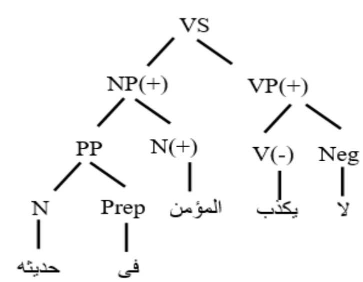


Fig. 2b. The Dependency Tree of a Negated Verb

Example 3: Noun Sentence
"الرجل غير كريم فى تصرفه"
NS:- NP NP
NP:- N
NP:- Neg NP
NP:- N PP
PP: - Prep N

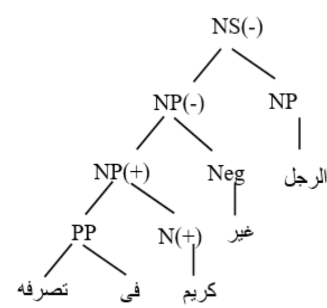


Fig. 2c. The Dependency Tree of a Noun Sentence

Example 4: Negated Noun Sentence
"المنافق ليس صادقا و لا امينا"
NS: - NP NP comp
NP: -N
NP: - N
Comp: - tool NP

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

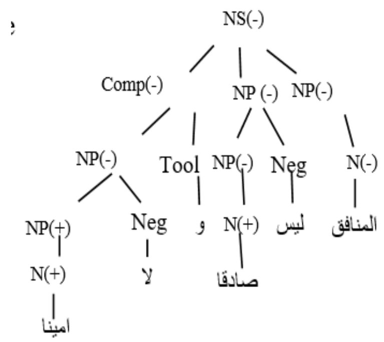


Fig. 2d. The Dependency Tree of a Negated Noun Sentence

Global Grammar
S : - NS | VS PP :-
Prep N
NS : - N NP | NP NP Comp
:- tool NP
NP : - N NP :-
N PP
NP : - Neg NP VP :-
V
VS : - VP NP VP :-
Neg V

4. IMPLEMENTATION WORK

To evaluate the performance of the adopted sentiment analysis model, it should be operated and tested. The model was implemented using Python and some natural language utilities that were used for some preprocessing operations such as tokenization, normalization, stopwords' rejection, and stemming. The model was run on a labtop with 4-GB memory size and 2.50 GHZ processing speed. The model also was run and tested using three test-beds or datasets as mentioned before. A brief descriptive statistic of the adopted datasets is mentioned as shown in Table IV.

Table 4: Descriptive Statistics of the Adopted Datasets

Statistical Items	**Arabic Style**		
	MSA	**Colloquial**	**Mixed**
Number of comments	340	1402	1742
Minimum number of tokens/comment	3	2	2
Maximum number of tokens/comment	870	159	870
Average number of tokens/comment	71	17	28
Number of tokens	24268	24272	48540

Some measurable criteria are used to evaluate the performance of the sentiment model. The evaluation criteria are: precision, recall, accuracy and error-rate and they are defined as follows:

$$\text{Accuracy} = \frac{\text{Tp} + \text{Tn}}{\text{Tp} + \text{Fp} + \text{Tn} + \text{Tp}} \tag{7}$$

$$\text{Error} - \text{rate} = \frac{\text{Fp} + \text{Fn}}{\text{Tp} + \text{Fp} + \text{Tn} + \text{Tp}} \tag{8}$$

(9) $$\text{Precision} = \frac{\text{Tp}}{\text{Tp+Fp}}$$

(10) $$\text{Recall} = \frac{\text{Tp}}{\text{Tp+Fn}}$$

Where T_p is the number of true positive reviews correctly classified, T_n is the number of true negative reviews correctly classified, F_p is the number of false positive reviews incorrectly classified as positive, F_n is the number of false negative reviews incorrectly classified as negative [8]. Moreover, some experiments were done and implemented. In each experiment, the above mentioned measurable criteria were calculated. The first group of experiments was done to evaluate the effect of using the general stemming and the light stemming. The second group of experiments was focused to measure the difference in performance using that adopted negation method and the proposed one. The third group of experiments was run to check the effect of using the important themes such as negations, intensifiers, and emotions. Such themes were conducted individually and compared with the original sentiment model (i.e without negation, intensifier, and emotions). The next experiments studied the effect of two-pairs of such themes with the original model. The effect of using negations and intensifiers, negations and emotions, intensifiers and emotions are also evaluated. The last experiments were run to see the effect of using all themes together with respect to the original one. It is important to mention that the effect of using negations, intensifiers, and emotions were developed and run using the light stemmer due to its better performance compared to the general stemmer. Figures 3-9 show the obtained results of the experiments (precision, recall, accuracy, and error) using the chosen datasets. Figure 3 shows the performance of the sentiment model using the general and light stemmers. Figures 4-6 show the behavior of the sentiment model with and without emotions, negations, and intensifiers respectively. Figures 7-9 on the other hand present respectively the model performance using negations (N), intensifiers (I), and emotions (E) in an individual form and combined form as well.

Moreover, the accuracy values of the predicted positive comments (as shown in Figure 10) were: 98.2%, 91.8%, and 85.5% for the Arabic style: MSA, mixed, and informal (colloquial), while the accuracy values for the predicted negative comments were respectively 93.2%, 92.6%, and 70.1% for the different Arabic styles.

5. DISCUSSION OF RESULTS

From the experimental work, it was noticed that: the sentiment model behaves better using the light stemmer than the using the general one. This is clean from the obtained results for the measuring criteria: accuracy, precision, recall, and error-rate. The behavior of the sentiment model was improved when adopting the emotions in the users' comments. i.e the emotions' effect improved the performance up to 11.2%, 13.4%, 12.1% and 8.3% compared with that behavior without emotions, for precision, recall, accuracy, and error-rate respectively. By considering the Arabic negation words in the comments, the model performance was improved by 16.1%, 14.2%, 15.1%, and 10.2% respectively for precision, recall, accuracy, and error-rate. The Arabic intensifiers also have a significant effect on sentiment analysis model. The performance was improved by about 9.1%, 11.2%, 10.2%and 7.3% respectively for precision, recall, accuracy, and error-rate. Moreover, the group of experiments was done to study the effect of adopting negation-intensifiers, negation-emotions, emotions-intensifiers on the users' comments. Using both the Arabic words (negations and emotions), the performance was improved by about 27.3%, 26.2%, 28.1%, and 18.4% respectively for precision, recall, accuracy, and error-rate. The model performance; on the other hand; was improved by about 24.2%, 22.6%, 25.2%, and 16.1% respectively for precision, recall, accuracy, and error-rate when the model considered both the negations and intensifiers words. It is also important to mention that the emotions and intensifies have a significant effect on the model performance. The model behavior was improved by 21.9%, 23.1%, 23.1%, and 14.1% for precision, recall, accuracy, and error-rate. The performance was also improved by about 20.2%, 18.8%, and 8.2% respectively for precision, recall, and error when adopting all themes compared to that performance without using any of them (i.e negations, intensifiers, and emotions). Moreover, the accuracy values of the predicted positive comments were 98.2%, 91.8%, and 85.8% while those values for predicting the negative comments were 93.2%, 92.6%, and 70.1% respectively for the comments written in MSA, Mixed Arabic, and informal Arabic.

6. CONCLUDING REMARKS

The web technology plays an important role in our life. The web contains a huge amount of information concerning the people's sentiment and opinions. This research work discussed a sentiment analysis model to analyze the people's opinions and emotions towards some comments collected from the Facebook. The collected comments were written in Arabic as the Arabic language in both challenges and intensity. The work analyzed the impact of preprocessing operations like noise rejection, stemming, and normalizations in the user's comments. The study also investigated some key themes that have a significant effect on the polarity of the sentence either positively or negatively. The adopted themes are: negation words, emotions, and intensifiers words. Several experiments were presented to see the effect of such themes either individually or combined. The performance of the sentiment analysis model was improved when adopting such themes. The best performance for the sentiment analysis model was achieved when using, negations, and intensifiers. Following this, the performance was also improved when using negations-emotions, then negations-intensifiers, then emotions-intensifiers respectively. Moreover, the accuracy values of the predicted positive comments were 98.2%, 91.8%, and 85.8% while the accuracy values for the predicted negative comments were 93.2%, 92.6%, and 70.1% when using comments written in MSA, Mixed Arabic, and informal Arabic (colloquial) respectively. Finally, the types of Arabic style used in the comments are

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

also significant. The best performance was for those comments written in MSA while the worst one for those adopting the informal Arabic style.

REFERENCES

[1] Osamah Ali Mohammed Ghaleb, and Anna Saro Vijendran, "Survey and Analysis of Recent Sentiment Analysis Schemes Relating to Social Media", The Indian Journal of Science and Technology, Vol.9, No.41, PP.1-16, November 2016.

[2] Salha Al-Osaimi, and Muhammad Badruddin Khan, "Sentiment Analysis Challenges of Informal Arabic Language", The International Journal of Advanced Computer Science and Applications, Vol. 8, No. 2, PP. 278-284, 2017.

[3] Ghahdeer AL-Sukkar, Ibrahim Aljarah, and Hamed Alsawalqah, "Enhancing the Arabic Sentiment Analysis Using Different Preprocessing Operators", The Proceedings of the New Trends in Information Technology NTIT, PP. 1-5, The University of Jordan, Amman, Jordan, April 25-27, 2017.

[4] I. Mohan, K, Janani, and M. Karthiga, "A Survey on Sentiment Analysis on Social Network Data", The International Journal of Scientific Research in Computer Science, Engineering and Information Technology, Vol.2, Issue 2, PP. 562-568, March-April 2017.

[5] Tajinder Singh, and Madhu Kumari, "Role of Text Pre-Processing in Twitter Sentiment Analysis", The 12th International Multi-Conference on Information Processing, ScienceDirect, Vol89, PP. 549-554, 2016.

[6] Hady ElSahar and Samhaa R. El-Beltagy, "Building Large Arabic Multi-domain Resources for Sentiment Analysis", Springer International Publishing Switzerland, Part II, LNCS, PP. 23-34, 2015.

[7] Ali Hamdi, Khaled Shalan and Anazida Zainal, "A Review on Challenging Issues in Arabic Sentiment Analysis", Journal of Computer Sciences, Science publication, Vol.12, No.9, PP. 471-481, 2016.

[8] Muhammad Zubair Asghar, Aurangzeb Khan, Shakeel Ahmad, Maria Qasim, and Imran Ali Khan, "Lexicon-enhanced Sentiment Analysis Framework using Rule-Based Classification Scheme", Journal PLOS ONE, PP. 1-22, February 23, 2017.

[9] Mahmoud Nabil, Mohamed Aly, and Amir F.Atiya, "ASTD: Arabic Sentiment Tweets Dataset", The Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, PP. 2515-2519, Lisbon, Portugal, September 17-21, 2015.

[10] Alaa Hamouda and Mohamed Rohaim, "Reviews Classification Using SentiWordNet Lexicon", The Online Journal on Computer Science and Information Technology (OJCSIT), Vol. 2, No. 1, PP. 120-123, January 2012.

[11] Xing Fang and Justin Zhan, "Sentiment Analysis using Product Review Data", The Journal of Big Data, Springer, Vol. 2, No.5, PP. 1-14, 2015.

[12] Anna Jurek, Maurice D. Mulvenna and Yaxin Bi, "Improved lexicon-based Sentiment Analysis for Social Media Analytics", The Securing Informatics, Springer, Vol. 4, No.9, PP. 1-13, 2015.

[13] Amira Shoukry, Ahmed Rafea , "Preprocessing Egyptian Dialect Tweets for Sentiment Mining", An Outcome Research Paper from the Project Entitled "Semantic Analysis and Opinion Mining for Arabic Web" the Project was Sponsored by ITIDA and Egyptian Industrial company LINK Department, PP. 47-56, 2015.

[14] Ahmed Y. Al-Obaidi, Venus W. Samawi, "Opinion Mining: Analysis of Comments Written in Arabic Colloquial", An Online Reference Downloaded from https://www.researchgate.net/publication/309678051_Opinion_Mining_ Analysis_of_Comments_Written_in_Arabic_ Colloquial, 2017.

[15] Khan F. H., Bashir S, and Qamar U. Tom, "Twitter Opinion Mining Framework using Hybrid Classification Scheme", Decision Support Systems, 57, PP. 245-257, 2014.

[16] https://github.com/mahmoudnabil/labr

[17] M. Z. Asghar, A. Khan, S. Ahmad, M. Qasim, I. A. Khan, "Lexicon-Enhanced Sentiment Analysis Framework Using Rule-Based Classification Scheme", The PLOS ONE Journal, PP.1-22 February 23, 2017.

[18] M. Nabil, M. Aly, A. Atiya, "ASTD: Arabic Sentiment Tweets Dataset", the Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing", Lisbon, Portugal, PP. 2515-2519, 17-21 September 2015.

[19] Alaa Hamouda, "Reviews Classification Using SentiWordNet Lexicon", The Online Journal on Computer Science and Information Technology (OJCSIT), Vol. 2, No. 1, January 2012.

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

[20] S. R. El-Beltagy, "NileULex: A Phrase and Word Level Sentiment Lexicon for Egyptian and Modern Standard Arabic", The Conference on Language Resources and Evaluation (LREC), Portorož, Slovenia, 2016.

[21] X. Fang and J. Zhan, "Sentiment Analysis using Product Review Data", The Journal of Big Data, Vol. 2, No.5, PP. 1-14, 2015.

[22] A. Jurek, Maurice 1D. Mulvenna and Yaxin Bi, "Improved Lexicon-based Sentiment Analysis for Social Media Analytics", The Secure Inform Journal, Vol. 4, No.9, 2015.

[23] Ammna Asmi and Tako Ishaya, "Negation Idetification and Calculation in Sentiment Analysis", The Second International Conference in Information Mining and Management, IMMM, PP.1-7, 2012.

[24] Arturo Monteyo-Ráez, Eugenio Martínez-Cámara, M.Tresa Martín-Valdivia, and L.Alfonso Ureña-López, "Ranked WordNet Graph for Sentiment Polarity Classification in twitter", Computer Speech and Language, Vol. 28, PP.93-107, 2014.

[25] Cambria Erik, Havasi Catherine, Hussain Amir, "SenticNet2: A Semantic and Affective Resource for Opinion Mining and Sentiment Analysis", The Proceedings of the Twenty-Fifth International Florida Artificial Intelligence Research Soceity Conference, USA, PP.202-207, 207, 2012.

[26] Anna Asmi and Tanks Ishaya, "Negation Identification and Calculation in Sentiment Analysis", The 2nd International Conference on Advances in Information Mining and Management (IMMM 2012), ISBN: 978-1-61208-227-1, PP. 1-7, 2012.

[27] Nádia F.F.da Silvaa Eduardo R. Hruschka and Estevam R. Hruschka Jr., "Tweet Sentiment Analysis with Classifier Ensembles", Decision Support Systems, Vol. 66, PP. 170-179, 2014.

[28] Abinash Tripathy, Ankit Agrawal, Santanu Kumar Rath, "Classification of Sentiment Reviews Using n-gram Machine Learning", Expert Systems with Applications, Vol. 57, PP.11-126, 2016.

[29] Umar Farooq, Hasan Mansoor, Antoine Nongaillard, Yacine Ouzrout, Muhammad Abdul Qadir, "Negation Handling in Sentiment Analysis at Sentence Level", The Journal of Computers, Vol. 12, No. 5, PP. 470-478, September 2017.

[30] Basant Agarwal and Namita Mittal, "Prominent Feature Extraction for Sentiment Analysis", Socio-Affective Computing Publishing, 2016.

[31] Buddhika H. Kasthuriarachchy, Kasun deZoysa, and H.L. Premarathne, "Context-Aware Sentiment Classification", The Fifteenth International Conference on Advances on ICT for Emerging Regions (ICTER 2015), Colombo, Sri Lanka, 2015.

[32] Anna Jurek, Yaxin Bi, Maurice Mulvenna, "Twitter Sentiment Analysis for Security-Related Information Gathering", The Proceedings of the 2014th IEEE Joint Intelligent and Security Informatics Conference (JISIC 2014), United Kingdom, PP. 48-55, 2014.

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

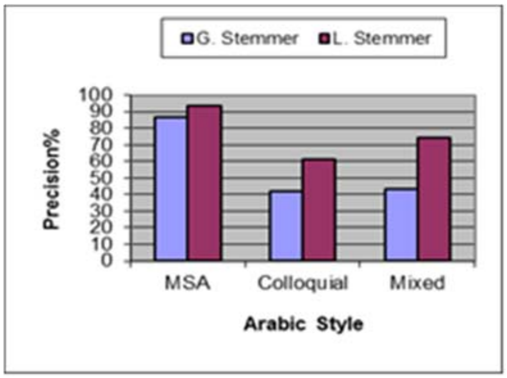


Figure 3a: Precison %

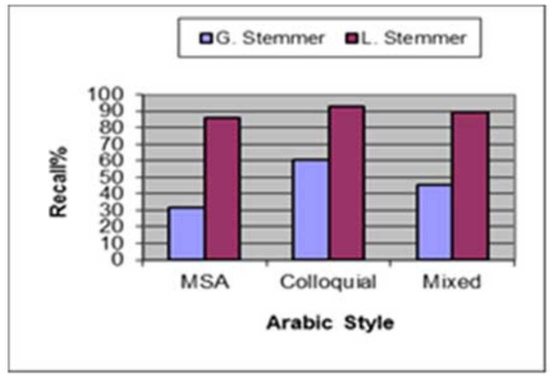


Figure 3b: Recal %

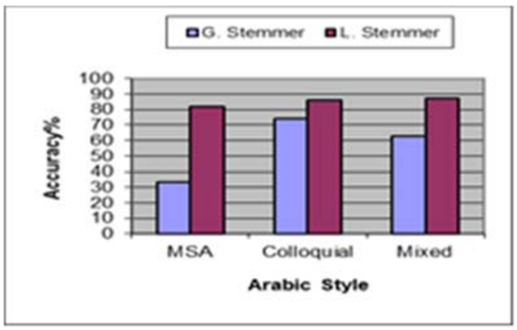
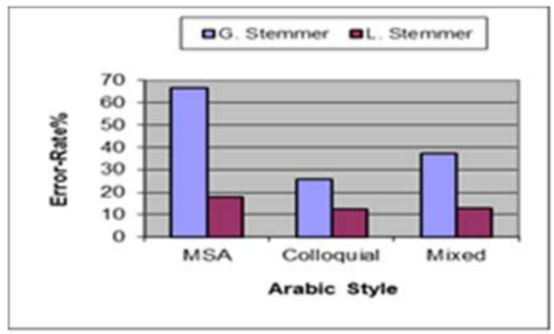


Figure 3c: Accuracy %



3d: Error-Rate %

Figure 3: Performance using General and Light Stemmer

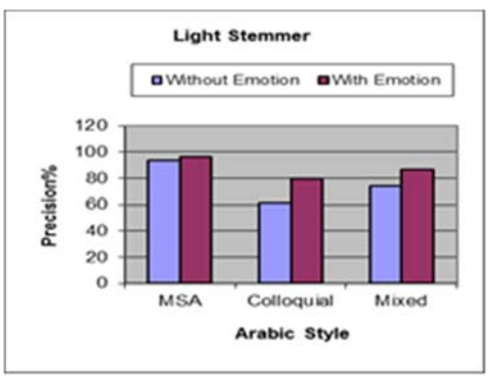


Figure 4a: Precison %

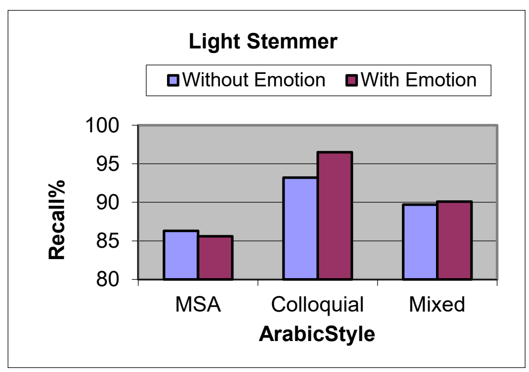


Figure 4b: Recal %

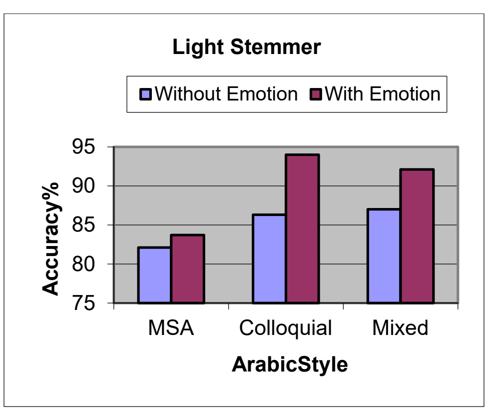


Figure 4c: Accuracy %

Figure 4: Performance of the Sentiment Model with and without Emotions

Journal of Theoretical and Applied Information Technology
15th November 2018. Vol.96. No 21

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

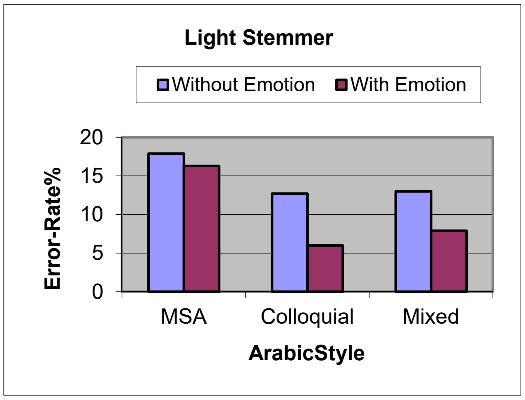


Figure 4d: Error-Rate %

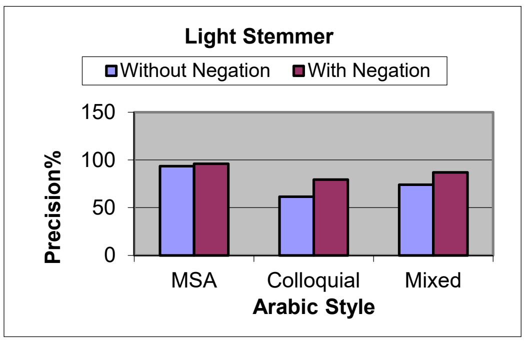


Figure 5a: Precison %

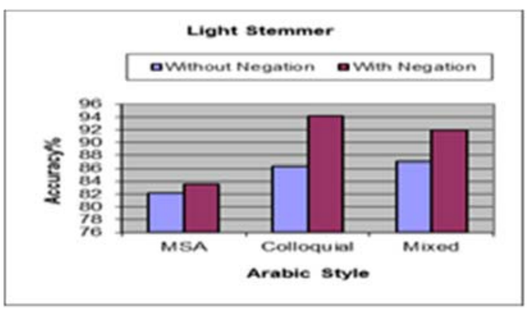


Figure 5c: Accuracy %

Figure 5: Performance of the Sentiment Model with and without Negations

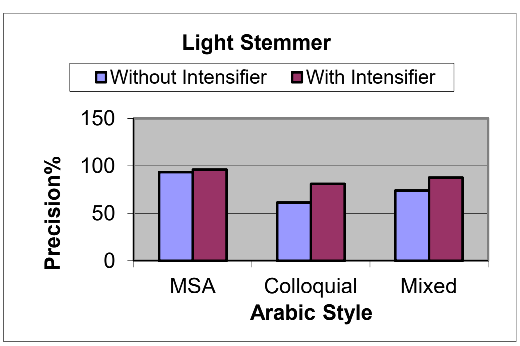


Figure 6a: Precison %

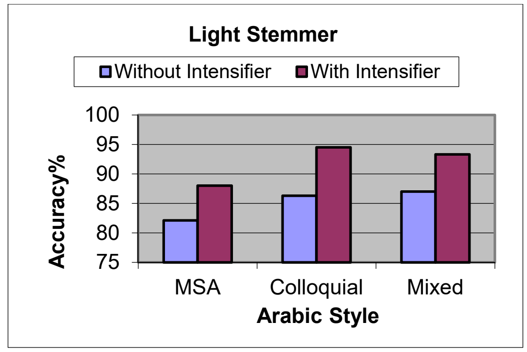


Figure 6c: Accuracy %

Figure 6: Performance of the Sentiment Model with and without Intensifiers

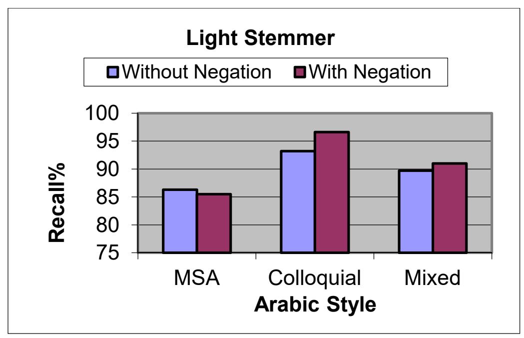


Figure 5b: Recal %

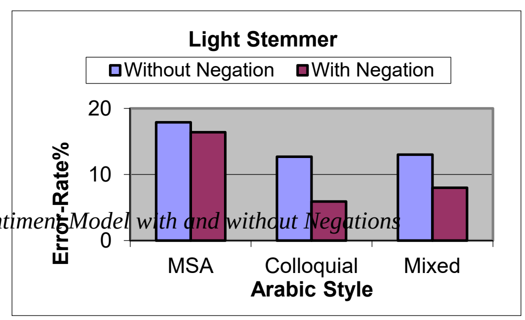


Figure 5d: Error-Rate %

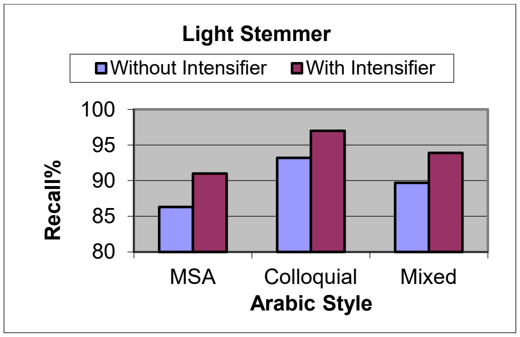


Figure 6b: Recal %

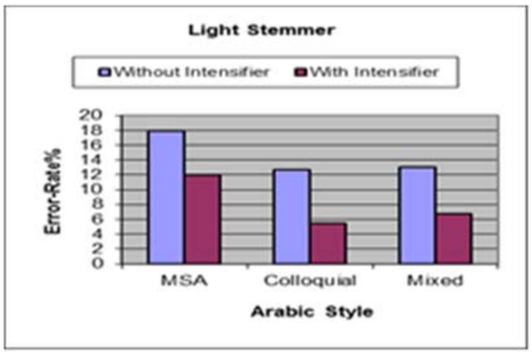


Figure 6d: Error-Rate %

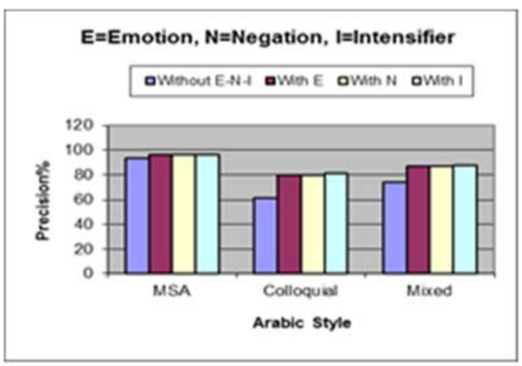


Figure 7a: Precison %

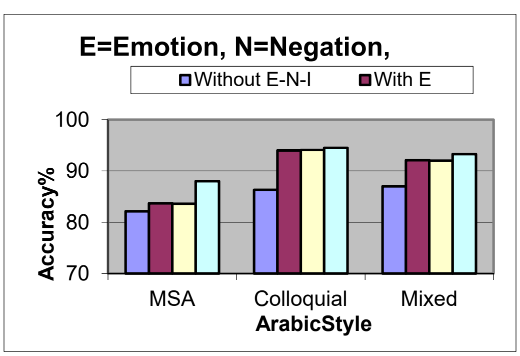


Figure 7c: Accuracy %

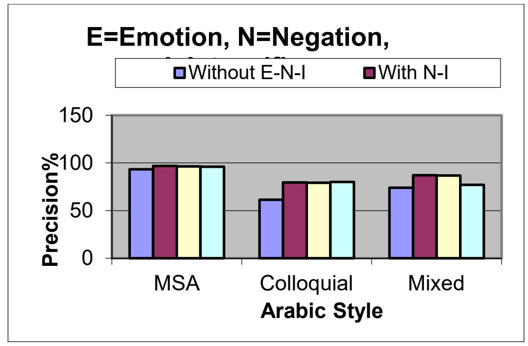


Figure 8a: Precison %

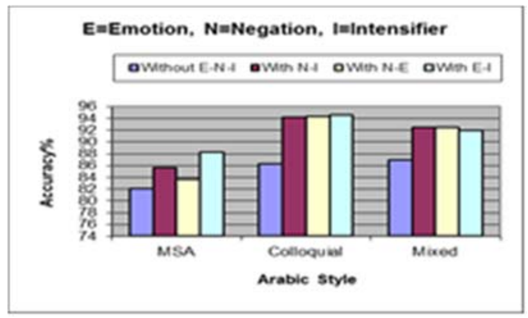


Figure 8c: Accuracy %

Figure 8: Performance of the Sentiment Model with Negations-Intensifires, Negations-Emotions, Emotions-Intensifiers and without

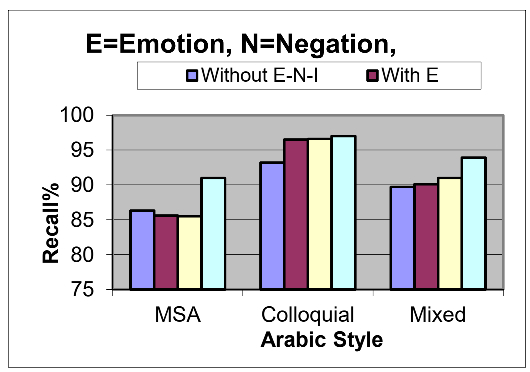


Figure 7b: Recal %

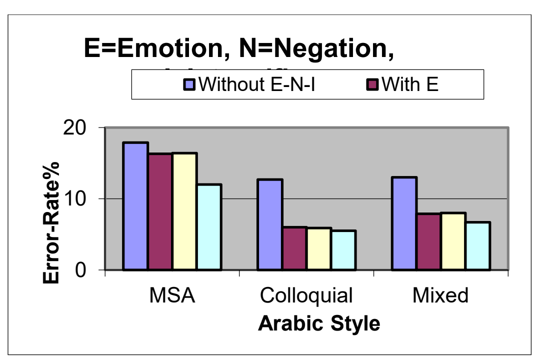


Figure 7d: Error-Rate %

Figure 7: Performane of the Sentiment Model with and without Emotions, Negations, and Intensifires

Journal of Theoretical and Applied Information Technology
15th November 2018. Vol.96. No 21

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

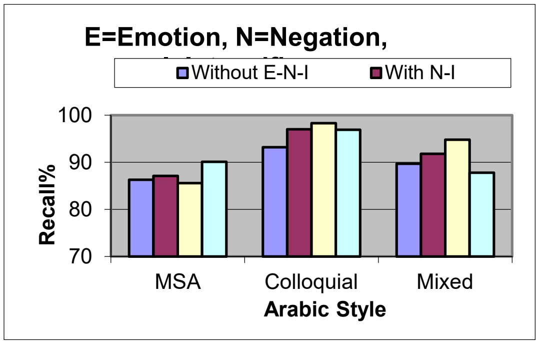


Figure 8b: Recal %

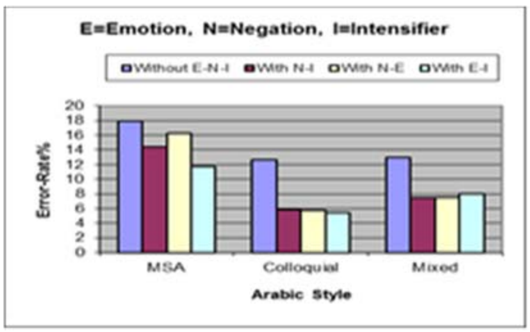


Figure 8d: Error-Rate %

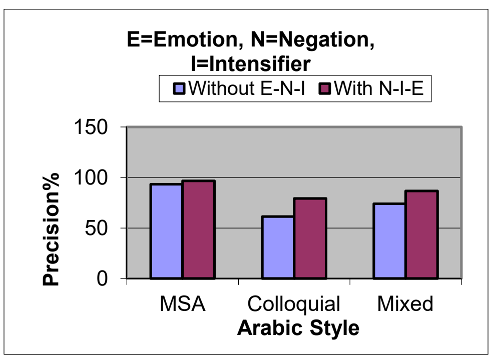


Figure 9a: Precison

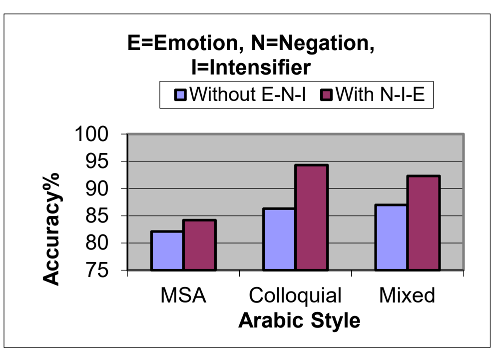


Figure 9c: Accuracy %

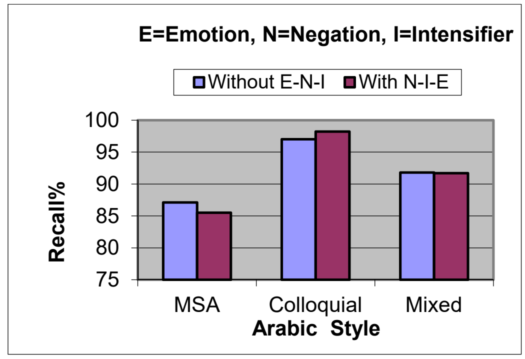


Figure 9b: Recal %

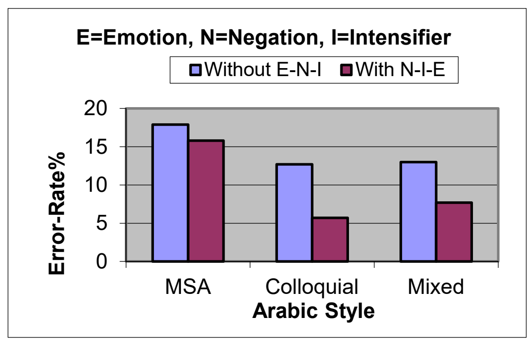


Figure 9d: Error-Rate %

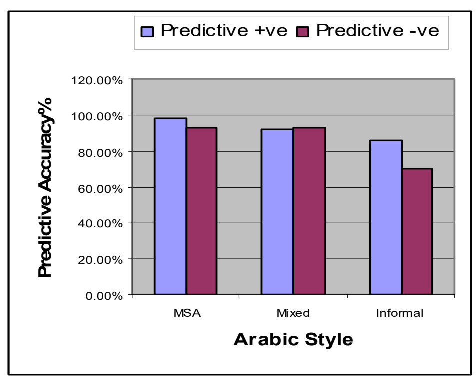


Figure 10: Predictive Accuracy for +ve and -ve Comments

Figure 9: Performance of the Sentiment Model with and without any of them: Negations, Intenmsifires, and Emotions