Sentiment Analysis for Arabizi Text

Rehab M. Duwairi
Department of Computer Information Systems,
Jordan University of Science and Technology,
Irbid 22110, Jordan
rehab@just.edu.jo

Mosab Alfaqeh, Mohammad Wardat, Areen Alrabadi
Department of Computer Science,
Jordan University of Science and Technology
Irbid 22110, Jordan

***Abstract*—This paper has used supervised learning to assign sentiment or polarity labels to tweets written in Arabizi. Arabizi is a form of writing Arabic text which relies on using Latin letters rather than Arabic letters. This form of writing is common with the Arab youth. A rule-based converter was designed and applied on the tweets to convert them from Arabizi to Arabic. Subsequently, the resultant tweets were annotated with their respective sentiment labels using crowdsourcing. This ArabiziDataset consists of 3206 tweets.**

Results obtained by this work reveal that SVM accuracies are higher than Naive Bayes accuracies. Secondly, removal of stopwords and mapping emoticons to their corresponding words did not greatly improve the accuracies for Arabizi data. Thirdly, eliminating neutral tweets at early stage in the classification improves Precision for both Naive Bayes and SVM. However, Recall values fluctuated, sometimes they got improved; on other times they did not improve

Keywords*— *Sentiment analysis; polarity classification; supervised learning; Arabic tweets; Arabizi

I. INTRODUCTION

Social media are popular channels for sharing ideas, opinions, viewpoints and sentiments. The Arab world's current circumstances have provided bloggers and commenters with numerous subjects to address [8, 21]. This huge content has encouraged researchers to address social media in their research. It is troublesome to manually process the continuous streams of social media content. Alternatively, researches used automated or semi-automated analysis which can handle such enormous amounts of social content effectively and efficiently.

Twitter enables users to send and read short messages called tweets [31]. It is free to use and its users can express their opinions openly which makes these tweets an ideal target for research.

Sentiment analysis (SA) or opinion mining is concerned with determining if a given review or comment is positive, negative or neutral. This has many benefits for businesses, education, commerce, health and many more. Unsupervised learning has been used for SA such as the work of Turney [30]. By comparison, supervised learning was used for SA such as the work of Pang et al [27]. Both works have addressed SA at the whole document level.

SA can be carried out at a finer level where the opinions portrayed by features or aspects mentioned in the tweets or reviews are extracted [10, 22, 25, 33].

Several SA methods rely on words' prior polarity to decide polarity labels of reviews or comments [9, 13, 16, 23, 24]. Also, machine learning, classification in particular, has been applied for SA as well [1, 3, 4, 5, 7, 14, 15, 17, 18].

In this work, we have used Naïve Bayes (NB) and SVM classifiers to assign polarity labels to tweets written in Arabizi after translating them into Arabic. Arabizi is a form of chat language where mainly the Latin alphabet is used to write Arabic words. A dataset, called the ArabiziDataset, which consists of 3206 tweets written using Arabizi and distributed over the Positive, Negative and Neutral classes, was collected, translated and annotated for the purpose of this research. A tailored built-in crowdsourcing tool was used to annotate the tweets with their respective polarity. Thorough experimentations were carried out to determine the effects of classifiers on the accuracy, the effects of preprocessing or filtering on the accuracy and the effects of removing the Neutral class from the dataset at an early stage and focusing on subjective tweets only. A rule-based converter was designed in this research to translate Arabizi tweets into Arabic ones.

The reported research has several novel contributions which include handling Arabizi tweets; designing the rule-based converter; and providing a deep and through analysis of SA for the Arabic language.

This paper is organized as follows. Section I provided an introduction to this research. Section II, by comparison, provides a list of related work. Section III explains, in details, our suggest framework for SA for the Arabic language. Section IV, summarizes the experiments and analyzes the results. Finally, Section V summarizes the findings of this paper and highlights future work.

II. BACKGROUND AND RELATED WORK

SA for Arabic is gaining popularity among researchers. Arabic is a morphologically rich language and this poses challenges for researchers working on Natural Language Processing (NLP), text mining or machine learning [19]. The peculiarities of the Arabic language which are related to SA are described in [11]. The remaining of this section lists related work which targets SA for the Arabic language.

978-1-4673-8614-2/16/$31.00 ©2016 IEEE

Abbasi et al [1] analyzes blogs written either in Arabic or English. The purpose of their study was to detect hostility on the web forums. Syntactic (Such as POS tags and stemming) and stylistic (Such as character n-grams) features were used to represent the comments. They conclude that combining syntactic and stylistic features improves the accuracy of their system.

Farra et al [20] relies on a set of features to represent each sentence. Examples of these features include frequency of positive, negative and neutral words in every sentence, frequency of negations, use of special characters, frequency of emphasis words, and frequency of contradiction words. The feature vectors of the documents were fed to the J48 Decision Tree classifier and the accuracy of classification was 62%.

Rushdi-Saleh et al. [28] applied the SVM and NB classifiers on a dataset that consists of 500 movie reviews written in Arabic. They used TF-IDF to weight the tokens of the reviews and they also stem the tokens. The accuracy of the SVM accuracy was 90% and the accuracy of the NB classifier was 84%

SAMAR [2] is a two-stage classifier which first distinguishes subjective sentences from objective ones. Subsequently, it determines the polarity of subjective sentences. The SVM light classifier was used for both stages. The dataset that they have experimented with consists of 8940 sentences.

Mourad and Darwish [26] used the NB classifier for analyzing 2300 tweets. A sentiment lexicon was also used to enhance the set of features that the NB classifier would use. The accuracy of deciding whether a tweet is subjective or objective, was equal to 76.6%. The accuracy of polarity detection was equal to 80.5.

Shoukry and Refae in [29] experimented with a dataset that consists of a 1000 tweets uniformly distributed between the positive and negative classes only. The SVM and NB classifiers were used. The results show that SVM outperformed NB with accuracy equals 72.6%.

The reported work in this paper is similar to few of the above research in the sense that we have used well-known classifiers for SA. However, our work has many novel aspects itself and adds value to the research on SA. For example, we relied on crowdsourcing to annotate the datasets. We provided a novel approach that deals with Arabizi tweets by developing the rule-based converter. Finally, we ran a large number of experiments that provide deep understanding of SA in Arabic.

III. THE FRAMEWORK OF TEXT POLARITY DETECTION

This section describes the main components of our suggested framework for determining the polarity of tweets written in Arabizi.

A. *Overall Architecture*

Our suggested framework for SA consists of the following stages:

- Collecting the dataset (Raw Data). Section III.B explains the methods used in collecting the dataset addressed by this work.
- Preprocessing the tweets – Every tweet was tokenized into words, also every emoticon in the tweet was mapped into its corresponding word, stopwords were also removed from that tweet, and lastly the weight of every token was calculated using the Binary Model. This means that 1 is assigned as a weight if the word is present in the current tweet regardless of its frequency; or 0 is used as a weight if the word is absent from the current tweet. Table I shows a sample of emoticons and shortcuts alongside their corresponding words in Arabic and their weights (i.e. sentiments); 1 means positive, -1 means negative and 0 means neutral. Preprocessing also includes converting tweets written in Arabizi into tweets written in Arabic using our built-in rule-based converter. The Arabizi converter is described in Section III.C and it is one of the major contributions of the current work. The output of the preprocessing stage is a cleaned dataset. This dataset is ready for labelling. Crowdsourcing was used to label the dataset.
- Classification stage, the NB and SVM classifiers were used to test the viability of the suggested framework. A detailed description of the various settings which were used with these two classifiers is found in Section IV (Experimentation and Result Analysis).

TABLE I. A SAMPLE OF EMOTICONS, THEIR CORRESPONDING ARABIC WORDS AND POLARITY LABEL

Emoticon/ Shortcut	Corresponding Arabic Word	Label/ Weight
:'(	بكاء	-1
O:)	وجهه ملائكي(بريئ)	1
3:)	وجهه شيطاني	-1
>:(	وجهه غاضب	-1
^_^	وجهه سعيد جدا	1
o.o	وجهه مرتبك	-1
:/	زعلان	-1
:)	سعيد	1
LOL	سعيد جدا	1
Isa	انشالله	0
Imo	برأيي	0
Tyt	اهلا بعودتك	1
Jak	جزاك الله خير	1

B. *Dataset Collection and Labelling*

The dataset, which was used in this research, was collected and annotated using Twitter API [32] and the Crowdsourcing Tool described in [12], respectively. This dataset is called the ArabiziDataset. It consists of tweets written in Arabizi [6]. Arabizi is a slang language used by Arab commenters where they use Latin letters to write Arabic words such as "mar7aba" which corresponds to "مرحبا" and means "Hello" in English. This dataset consists of 3206 tweets. However, before rating these tweets they were converted from Arabizi to Arabic using our built-in, rule-based converter (Refer to Section III.C). After conversion, the tweets were labelled using the crowdsourcing tool described in [12]. The crowdsourcing tool prompts users with tweets, one tweet at a time, the viewer will choose a label out of the following labels: Positive, Negative, Neutral and Not Applicable. Tweets with "Not Applicable" label are removed from the dataset as these are usually empty tweets or contain missed up text and tags. The tweets and their corresponding labels are stored in a database. Every tweet is required to have at least three labels to be considered finished labelling. Majority voting was used to determine the final label. Table II shows the distribution of the tweets over the set of polarity labels (i.e. positive, negative or neutral). These human labels are considered the absolute golden truth and the accuracy of the classifiers is calculated with respect to these labels. The ArabiziDataset is not balanced – where 1803 tweets express positive sentiment, 831 tweets express negative sentiment and 572 tweets are neutral. Table III shows some statistics about the tweets in the dataset such as average number of words and average number of characters per tweet.

TABLE II. DISTRIBUTION OF TWEETS OVER POLARITY LABELS

Label	ArabiziDataset
Positive	1803
Negative	831
Neutral	572
Total	3206

TABLE III. STATISTICS ABOUT THE ARABIZI DATASETS

	Arabizi Dataset		
	Positive	Negative	Neutral
Number of tweets	1803	831	572
Average number of characters/tweet	57.7	65.9	42.4
Average number of words/tweet	13.15	13.8	8.5

C. Arabizi Converter

Arabic language has three varieties: Traditional Arabic, Modern Standard Arabic (MSA) and colloquial Arabic [19]. The first variety is mostly used in religious scripts; the second is used in formal meetings; and the third is dependent on the country where the dialect is spoken. Every Arab country has its own spoken dialect. In fact there are varieties in the dialects within the same country. Colloquial Arabic is mostly spoken not written. However, when considering comments or reviews on social medial channels, we notice that colloquial Arabic is spread and mixed with MSA. All the previous forms of Arabic are written using Arabic Alphabet; which consists of the following 28 characters in addition to Hamza:

ا ب ت ث ج ح خ د ذ ر ز س ش ص ض ط ظ ع غ ف ق ك ل م ن ه و ء

A new writing style has emerged in social medial for Arabic bloggers; this is called Chat language or Arabizi. Here the writers use Latin letters to write Arabic words. Every Arabic character is mapped into its corresponding Latin letter as shown in Table IV.

TABLE IV. ARABIC LETTERS AND THEIR CORRESPONDING LATIN LETTERS USED IN ARABIZI

Character in Arabic Language	Corresponding Character in Arabizi	Character in Arabic Language	Corresponding Character in Arabizi
ا	**a**	ط	**T**
ب	**B**	ظ	**6'**
ت	**T**	ع	**3**
ث	**t', th or 4**	غ	**3'**
ج	**j or g**	ف	**F**
ح	**7**	ق	**8**
خ	**5 ot 7'**	ك	**K**
د	**D**	ل	**L**
ذ	**d'**	م	**M**
ر	**R**	ن	**N**
ز	**Z**	ه	**H**
س	**S**	و	**w or o**
ش	**$ or sh**	ي	**e or i**
ص	**9**	ئ	**2**
ض	**9'**	ء	**2**

Comments or reviews written in Arabizi are usually considered spam by researchers on SA and thus are removed from the dataset. In this work, we have opted to use a different approach – where we have designed a rule-based converter that will map comments or tweets written in Arabizi to Arabic. The rule-based converter operates by firstly tokenizing a tweet into a sequence of words or tokens using white spaces and punctuation marks as separators. Secondly, every Arabizi word is converted into an Arabic word by mapping the Arabizi letters to their corresponding Arabic letters using the mappings specified in Table IV. Table V shows a few examples of comments written in Arabizi and their corresponding comments written in Arabic generated automatically by our converter. The results of the transformation are very good but are not free of errors. To the best of our knowledge, this work is the first to provide this treatment for Arabizi. As we have pointed above, Arabizi is typically removed from the comments.

IV. EXPERIMENTATION AND RESULT ANALYSIS

Table VI shows the macro-Precision and macro-Recall obtained when applying NB and SVM. The macro-Precision for the NB classifier was equal to 0.504 without filtering and was equal to 0.505 with filtering; this means that filtering hardly improves the Precision. The table also shows the macro-Precision for SVM; as it can be seen; Precision was not greatly improved when filtering was used. SVM gave slightly better results when compared with NB. As far as macro-Recall is concerned, Table VI illustrates that there were slight improvements when filtering is obtained. Generally, the values of Precision and Recall are low; less than 60%. Previous studies have shown that filtering such as removing stopwords and stemming generally gives accuracies approaching 90% when dealing with Arabic language [11]. This indicates that dealing with Arabizi not necessarily gives comparable results to Arabic. This may have several reasons: one of them is that the Arabizi converter does not give ideal results. That is not all Arabizi words are perfectly converted to their respective MSA counterparts.

TABLE V. ARABIZI TWEETS AFTER BEING CONVERTED TO ARABIC TWEETS USING THE ARABIZI CONVERTER

Tweets in Arabizi	Tweets converted to Arabic	Label
f3la 89h nja7 wf5r..	فعلا قصه نجاح وفخر	Positive
mwt fi 3'i6'k ,,,, wlm3lwmatk aih alsathj ,,,, 80% atha m$ ak4r mmn i7f6'wn al8ran hm mn alf8ra2 aih alsathj ,,,,, waln86h al4anih hi d3wh llf8ra2 wala3'nia2 l7f6' ktab allh 7ta itmkn mn aktsab hthh almn7h	موت في غيطغك ولمعلوماتك ايه الساثج ق اثا مش اكثر ممن يحفطغون القران هم من الفقراء ايه الساثج والنقطه الثانيه هي دعوه للفقراء والاغنياء لحفطغ كتاب الله حتا يتمكن من اكتساب هثه المنحه	Negativ e
asmi m$ mbin	اسمي مش مبين	Neutral
a$ alm6lwb nt$hwn i3ni wla $w	اش المطلوب نتشهون يعني ولا شو	Positive
an $a2allh dimaً b5ir ia ardn	ان شاءالله ديما بخير يا اردن	Positive
al7l alw7id 3إ'la8 aljam3h w3mrh la 7da t3lّm	الحل الوحيد غلاق الجامعه وعمره لا حدا تعل م	Negativ e
alsla7 alabi9' mjwd 3la alb9tat fi $ar3 aljam3h aljwab 3ndk	السلاح الابيض مجود علا البصتات في شارع الجامعه الجواب عندك	Negativ e

TABLE VI. PRECISION AND RECALL VALUES WITH AND WITHOUT FILTERING

	NB		SVM	
	Filter	NoFilter	Filter	NoFilter
Macro-Precision	0.505	0.504	0.555	0.549
Macro-Recall	0.555	0.537	0.587	0.584

Figure 1 illustrates the Precision values for the Positive, Negative and Neutral classes for the ArabiziDataset. We notice that the class Positive scored the highest Precision when compared to the other two classes. Recall that the ArabiziDataset is skewed in favor of the Positive class; the number of tweets which express positive sentiment equals to 1803 tweets. This skewedness helped the classifiers in building more accurate models for the Positive class. The Precision of the Positive class did not improve with filtering for both classifiers; in fact it became slightly lower. By comparison, the Precision of the Negative class was improved when the NB classifier was used but was slightly lowered when the SVM classifier was used. The Precision of the Neutral class was improved when filtering was applied using the SVM classifier.

Figure 2 depicts the Recall values for the three classes. As it can be seen, the Recall values improved for the Positive class when either the NB or the SVM classifiers were used with filtering. Recall becomes lower in the case of the Negative and Neutral classes. Also, Figure 2 shows that the values of Recall for both the Negative and Positive classes are low when they are compared with maximum possible Recall value of 1.0 or when compared with the Recall of the Positive class.

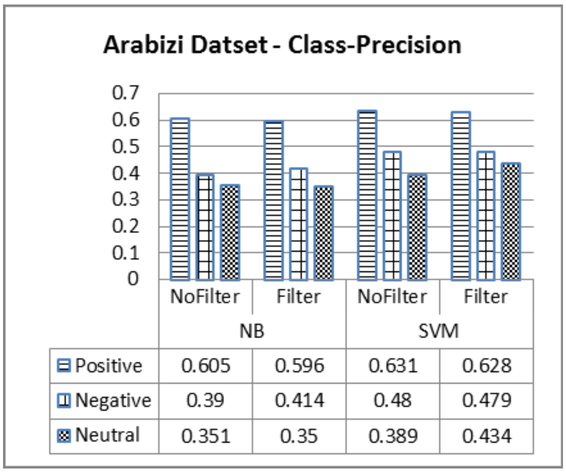

Fig. 1. Precision for the Positive, Negative and Neutral Classes

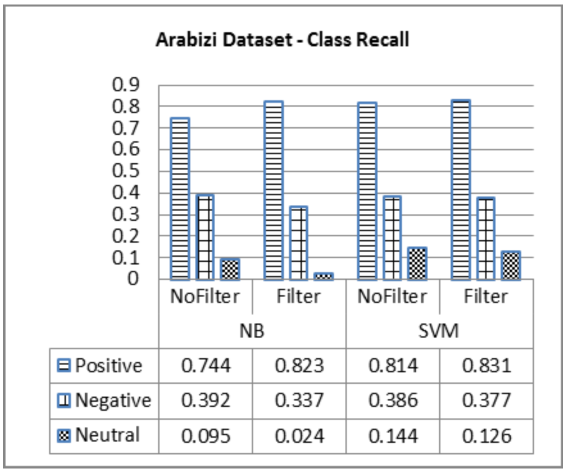

Fig. 2. Recall for the Positive, Negative and Neutral Classes

The findings depicted in Figure 1 and Figure 2 demonstrate that the NB and SVM classifiers did a very good job in predicting the correct examples of the class Positive (High Recall) and in eliminating the false examples/tweets of the Positive class (High Precision). The Precision and Recall of the Negative and Neutral classes were not good but the performance of the SVM classifier supersedes the performance of the NB classifier. The Recall values of the Neutral class were particularly very low which means that the classifiers did not do a good job in correctly classifying instances of the Neutral class. In other words, too many of the instances in the Neutral class were misclassified to belong to either the Negative class or the Positive class (Mostly to the Negative class). To overcome this problem, we have designed a new experiment, in which the classification is carried out in two stages. In the first stage, the classifier learns to distinguish objective (i.e. Neutral) tweets from subjective (i.e. positive or negative) tweets. In the second stage, the classifier learns to discriminate positive tweets from negative tweets. The results of this experiment are shown in Section IV.A.

A. Removing Objective Tweets from the Datasets

Table VII and Table VIII show the effects of removing neutral tweets from the ArabiziDataset on the Precision. The tables show that there are improvements especially for the Positive class.

TABLE VII. PRECISION VALUES OF THE NB CLASSIFIER WITH AND WITHOUT NEUTRAL TWEETS

	NB before removing the Neutral Class		NB After Removing the Neutral Class		Improvement	
	NoFilter	Filter	NoFilter	Filter	NoFilter	Filter
Positive Class	0.605	0.596	0.742	0.736	0.137	0.14
Negative Class	0.39	0.414	0.456	0.481	0.066	0.067

TABLE VIII. PRECISION VALUES OF THE SVM CLASSIFIER WITH AND WITHOUT NEUTRAL TWEETS

	SVM before removing the Neutral Class		SVM After Removing the Neutral Class		Improvement	
	NoFilter	Filter	NoFilter	Filter	NoFilter	Filter
Positive Class	0.631	0.628	0.758	0.749	0.127	0.121
Negative Class	0.48	0.479	0.574	0.564	0.094	0.085

Table IX and Table X illustrate the effects of removing the Neutral class from the ArabiziDataset on Recall. We notice that the improvement is marginal and in some cases removing the Neutral class adversely affected the Recall values.

V. CONCLUSIONS AND FUTURE WORK

This paper has presented a framework for dealing with sentiment analysis embedded in tweets written in Arabizi. A rule-based converter was designed and applied on the tweets. The resultant tweets, which are now written in Arabic, were annotated with their respective sentiment labels using crowdsourcing. A tool [12] was designed which allows volunteer users to register to carry out labelling. After registration, the users are presented with the tweets, one by one, and the users will choose the proper label. A tweet is considered labeled when three different users have assigned labels to that tweet. Majority voting was used to determine the final label. The ArabiziDataset consists of 3206 tweets. It is a skewed dataset in favor of the Positive class. To the best of our knowledge, the current work is the first to deal with tweets written in Arabizi for sentiment analysis and to develop a converter for this purpose.

TABLE IX. RECALL VALUES OF THE NB CLASSIFIER WITH AND WITHOUT NEUTRAL TWEETS

	NB before removing Neutral Class		NB After Removing Neutral Class		Improvement	
	NoFilter	Filter	NoFilter	Filter	NoFilter	Filter
Positive Class	0.744	0.823	0.77	0.821	0.026	-0.002
Negative Class	0.392	0.337	0.418	0.36	0.026	0.023

TABLE X. RECALL VALUES OF THE SVM CLASSIFIER WITH AND WITHOUT NEUTRAL TWEETS

	SVM before removing Neutral Class		SVM After Removing Neutral Class		Improvement	
	NoFilter	Filter	NoFilter	Filter	NoFilter	Filter
Positive Class	0.814	0.831	0.862	0.869	0.048	0.038
Negative Class	0.386	0.377	0.404	0.367	0.018	-0.01

Results obtained by this work reveal the following facts when dealing with the current ArabiziDataset. First, SVM results supersede NB results. The Precision and Recall given by SVM were higher than those given by NB. Secondly, Arabizi data were not greatly affected by filtering. Filtering, in the current work, removes stopwords and maps emoticons to their corresponding Arabic words. Thirdly, two-stage classification – where neutral tweets were removed at early stage, improves Precision for both NB and SVM. However, Recall values fluctuated, sometimes they got improved; on other times they did not improve.

ACKNOWLEDGMENT

This Project is supported by Jordan University of Science and Technology, grant No. 20130174.

REFERENCES

[1] A. Abbasi, C. Hsinchun, and S. Arab. Sentiment analysis in multiple languages: Feature selection for opinion classification in Web forums. ACM Transactions on Information Systems (TOIS). Vol. 26, no. 3, 2008, pp. 1-34.

[2] M. Abdul-Mageed, S. Kubler, and M. Diab. SAMAR: A system for subjectivity and sentiment analysis of Arabic Social Media. 3rd Workshop on Computational Approaches for Subjectivity and Sentiment Analysis WASSA, Satellite Workshop, Jeju, Koera, June, 2012.

[3] K. Ahmad, D. Cheng, and Y. Almas. Multi-lingual Sentiment Analysis of Financial News Streams. Proceedings of the 1st International Workshop on Grid Technology for Financial Modeling and Simulation, Palermo, Italy, 2006.

[4] K. Ahmad, and Y. Almas. Visualizing Sentiments in Financial Texts. Proceedings of the Ninth International Conference on Information Visualization. 2005, PP. 363 – 368, Washington, USA.

[5] G. Alec, R. Bhayani, and L. Huang. Twitter sentiment classification using distant supervision. CS224N Project Report, Stanford, 2009, pp. 1-12.

[6] http://en.wikipedia.org/wiki/Arabic_chat_alphabet. Last Accessed 26-Dec-2015.

[7] E. Boiy, and M.F. Moens. A Machine Leaning Approach to Sentiment Analysis in Multilingual Web Texts. Information Retrieval. Vol. 12, Issue 5, 2009, pp. 526 – 558.

[8] N. J. DeLong-Bas, The New Social Media and the Arab Spring" Oxford Islamic Studies Online, Oxford Press. http://www.oxfordislamicstudies.com/Public/focus/essay0611_social_media.html. Last accessed 26-Ndec-2015.

[9] K. Denecke. Are SentiWordNet Scores Suited for Multi-domain Sentiment Classification? In Fourth International Conference on Digital Information Management (ICDIM), 2009. 2009, pp. 33-38, Nov, Ann Arbor, MI.

[10] X. Ding, B. Liu, P.S. Yu, A holistic lexicon based approach to opinion mining. Proceedings of the International Conference on Web Search and Web Data Mining (WSDM'08), 2008, pp. 231–239.

[11] R. Duwairi, M. El-Orfali. A Study of the Effects of Preprocessing Strategies on Sentiment Analysis for Arabic Text, Journal of Information Science. Vol. 40(4), 2014, pp. 501-513.

[12] R. Duwairi, R. Marji, N. Sha'ban, S. Rushaidat. Sentiment Analysis in Arabic Tweets, Proceedings of the 5th International Conference on Information and Communication Systems, Irbid, Jordan, April 1-3, 2014.

[13] R. Duwairi, Mohammad A. Alshboul. Negation-aware Framework for Sentiment Analysis in Arabic Reviews, The 2nd International Workshop on Social Networks Analysis, Management and Security (SNAMS - 2015), August 2015, Rome, Italy.

[14] R. M. Duwairi. Sentiment Analysis for Dialectical Arabic. The 6th International Conference on Information and Communication Systems (ICICS), Amman, Jordan, 7-9 April, 2015.

[15] R. M. Duwairi, Islam Qarqaz, Arabic Sentiment Analysis using Supervised Classification, The 1st International Workshop on Social Networks Analysis, Management and Security (SNAMS - 2014), August 2014, Barcelona, Spain.

[16] R. M Duwairi, Nizar Mahyoub, Saleh Alrefai, Detecting Sentiment Embedded in Arabic Social Media – A Lexicon-based Approach. Journal of Intelligent and Fuzzy Systems, vol. 29, No. 1 pages: 107-117, 2015, I O S Press, Netherlands.

[17] A. El-Halees A. Arabic Opinion Mining Using Combined Classification Approach. Proceedings of the International Arab Conference on Information Technology (ACIT), 2011, Riyadh, Saudi Arabia, Dec.

[18] M. Elhawary, and M. Elfeky. Mining Arabic Business Reviews. Proceedings of the IEEE International Conference on Data Mining. 2010, pp 1108 – 1113, Dec, Mountain View, USA.

[19] A. Farghaly, and K. Shaalan. Arabic Natural Language Processing: Challenges and Solutions. ACM Transactions on Asian Languages Information Processing, Vol. 8, No. 4, Article 14, 2009.

[20] N. Farra, E. Challita, R.A. Assi, H. Hajj. Sentence-Level and Document-Level Sentiment Mining for Arabic Texts. Proceedings of the ICDM Workshop, 2010. pp. 1114-1119.

[21] P. N. Howard, A. Duffy, D. Freelon, M. Hussain, W. Mari, M. Mazaid, Opening Closed Regimes, What Was the Role of Social Media During the Arab Spring? http://pitpi.org/wp-content/uploads/2013/02/2011_Howard-Duffy-Freelon-Hussain-Mari-Mazaid_pITPI.pdf. Last accessed 26-Dec-2015.

[22] M. Hu, B. Liu, Mining opinion features in customer reviews. Proceedings of the Nineteenth National Conference on Artificial Intelligence (AAAI), 2004, pp. 755–760.

[23] K. Jaap, M.J. Marx, J. Robert, and M.D. Rijke. Using WordNet to measure semantic orientations of adjectives. Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC 2004). Vol. IV, 2004, pp. 1115–1118.

[24] M. Koppel, Schler J., The importance of neutral examples for learning sentiment. Computational Intelligence, 22(2), pp. 100-109.

[25] B. Liu, Sentiment Analysis and Opinion Mining, Morgan & Claypool Publishers, 2012.

[26] A. Mourad, and K. Darwish. Subjectivity and sentiment analysis of modern standard Arabic microblogs. Proceedings of the 4th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, 2013, pp. 55-64, Atlanta, Georgia, USA.

[27] B. Pang, L. Lee, and S. Vaithyanathan, Thumbs up? Sentiment classification using machine learning techniques. Proceedings of the 40th Annual Meeting on Association for Computational Linguistics (ACL 2002), 2002, pp. 79–86.

[28] M.R. Saleh, M.T. Martin-Valdivia, L.A. Urena-Lopez, and J.M. Perea-Orteg A. OCA: Opinion Corpus for Arabic. Journal of the American Society for Information Science and Technology. Vol. 62, Issue. 10, 2011, pp 2045-2054.

[29] A. Shoukry, and A. Rafea. Sentence-level Arabic sentiment analysis. Proceedings of Collaboration Technologies and Systems (CTS). 2012, pp. 546-550.

[30] P. Turney, Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. Proceedings of the 40th Annual Meeting on Association for Computational Linguistics (ACL 2002), 2002, pp. 417–424.

[31] https://twitter.com, last accessed 26-Dec-2015.

[32] Twitter API, https://dev.twitter.com/docs/api/1/get/search, last accessed 26-Dec-2015.

[33] M. Wang, G. Li, Z. Lu, Y. Gao, T. Chua, When Amazon Meets Google: product visualization by exploring multiple information sources, ACM Trans. Int. Technol. 12 (4) (2013) (Article 12).

[34] http://www.cs.waikato.ac.nz/ml/weka/. Last accessed 26-Dec-2015.