Journal of Theoretical and Applied Information Technology
20th October 2013. Vol. 56 No.2

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

OPINION ANALYSIS FOR TWITTER AND ARABIC TWEETS: A SYSTEMATIC LITERATURE REVIEW

[1]MNAHEL AHMED IBRAHIM, [2]NAOMIE SALIM

[1]PHD candidate., Department of Computer Science, collage of Computer Science & Infromation Technology,Sudan University Of Scince and Technoloy, Sudan

[2] Prof., Department of Information Systems, Faculty of Computer Science & Information Systems, Universiti Teknologi Malaysia, Malaysia

E-mail: [1]manahil_a@sustech.edu, [2]naomie@utm.my

ABSTRACT

The objective of this paper is to present the current evidence relative to twitter opinion mining in general and also the current state of Arabic tweets' opinion mining. The researcher performed a systematic literature review (SLR) to investigate features and methods used for twitter opinion mining and if those features and methods have been used for Arabic tweets opinion mining. Sixty five papers were used in our synthesis of evidences. Results showed that n-grams features are the most features used for twitter sentiments analysis and also for Arabic tweets. The most common methods used for twitter sentiments analysis is the Lexical based classification using Naive Bayes (NB) and Support Vector Machines (SVM), which are also used for Arabic tweets. In addition, evidence related to subjectivity and opinion target for twitter are highlighted. The results of this SLR show gaps in the research field: first, the lack of studies focusing on multilingual twitter sentiments analysis. Second, the lack of studies that investigate Arabic tweet opinion target. The third is the lack of studies investigating Arabic tweet subjectivity.

Keywords: *Opinion Analysis, Arabic Opinion Mining, Twitter, Systematic Review.*

1. INTRODUCTION

In general, the text generated by users can contain facts as well as opinions. Facts are objective expressions about entities, events and their attributes, e.g. ”I bought Sony camera yesterday” whereas opinions are subjective expressions of sentiments, feelings, attitudes, emotions, appraisals or opinions toward entities, events and their attributes, e.g. ”I really love Galaxy III” [1] and [2]. Sentiment analysis or opinion mining refers to the application of natural language processing, com- putational linguistics and text analysis to identify and extract subjective information in source materials [7]. It is the computational study of opinions, sentiments and emotions expressed in text on anything such as people, products or events, etc. [1].

Unfortunately, not all subjective sentences contain opinions, e.g. ”I want a screen with good resolution” and not all objective sentences contain no opinions, e.g. ”The G-Tide tablet screen broke in just two days!” Opinion mining consists of a set of techniques that are used in solving different types of problems, such as determining opinionated sentence which is a sentence that expresses opinions either explicit or implicit and determining its orientation or polarity which can be positive, negative or neutral. Any opinion has an opinion source or holder which is a person or organization that expresses the opinion [3], in most cases the author itself; and opinion target which is a topic on which opinion are expressed. There are two main types of opinions: the first one is regular opinions: opinion expressions on some target entities which can be either direct opinion, e.g. ”Galaxy II is so cool.” or indirect opinions, e.g. ”After taking the drug, my pain has gone”. And the second one is comparative opinions: comparisons of more than one entity. e.g., ”Galaxy II is better than iPhone 4.”[1] and [4].

Twitter is a real-time social network and micro blog service that enables users to send their opinions ,ideas and news in form of text message that called tweets which is 140 characters long [5] and [6]. Arabic is a Semitic language [9] and consists of different regional dialects, it is an official language of 22 countries worldwide and there are more than 350 million people spoken Arabic and it is the fastest-growing language on the web (with an annual growth rate of 2,501.2%

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

in the number of Internet users as of 2010, compared to 1,825.8% for Russian, 1,478.7% for Chinese and 301.4% for English) [8].

A. Sentiment Analysis For Twitter

The researcher applied a systematic literature review (Sec- tion II) in assessing existing twitter Sentiment analysis litera- ture on Sentiment analysis techniques for twitter and Arabic tweets. The researcher presents SLR results by integrating evidence into patterns that can be used to understand the current state of the art of research in twitter Sentiment analysis. The researcher believes that this can help researcher who want to perform Arabic Sentiment analysis for twitter in a better way. Additionally, findings from the analysis are presented and gaps in the existing body of knowledge are highlighted. These suggest key areas of focus for further research. Section II describes the method the researcher used in SLR. Section III reports the results of SLR based on the synthesis of evidence. Section IV presents a discussion of the key findings and future work. Section V presents conclusions from the review.

2. THE REVIEW METHOD

A. Introduction

An SLR is defined as a process of identifying, assessing, and interpreting all available research evidence with the purpose to provide answers for specific research questions [8]. It is a tool that aims to produce a scientific summary of the evidence in a particular area, in contrast to ”traditional” narrative review [9].

B. Research Questions

Table I shows the Population, Intervention, Comparison, Outcomes, and Context (PICOC) structure of the research questions. In SLR, the researcher included all studies that in- vestigated opinion mining of Twitter, regardless of whether or not concentrate on Arabic tweets. Considering the importance of understanding the factors to gain effective and efficient practice of opinion mining, a systematic review is needed to be held to assess the availability of existing research with regard to the issues of different types of techniques, features and corpus. It could be further suggest the gap or important areas of future studies. Therefore, this protocol is developed as a framework to conduct the systematic review in opinion mining focusing on opinion mining for Arabic tweets.

Table I: Suummary of Pico

Population	Tweets
Intervention and Comparison	Opinion mining of Arabic Tweet
Outcomes	Measurement of Precision and recall
Study Design	Comparison between techniques used in tweets and Arabic tweets
Context	Within the domain of opinion mining and information retrieval.

The aim of the proposed systematic review is to assist researchers in deciding which features and techniques are most highly regarded and critical for further exploration when conducting opinion mining for Arabic tweets. Therefore, the SLR aimed to answer the following primary research question (RQ):

1) Primary question: What are the techniques and features that can be used for Arabic opinion mining in tweet?

2) Secondary subquestions: Our SLR also aimed to answer the following secondary subQuestions:

1) What are the features and techniques that have been used for subjectivity identification of tweets? Are these features and techniques used for Arabic tweets?
2) What are the features and techniques that have been used for polarity identification of tweets? Are these features and techniques used for Arabic tweets?
3) What are the features and techniques that have been used for opinion target identification of tweets? Are these features and techniques used for Arabic tweets?
4) What are the fuzzy aspects that have been studied for twitter opinion mining? Are they used for Arabic tweets?
5) What are the Arabic corpora that have been used in Arabic opinion mining? Are they used for twitter?
6) What are the metrics that have been used to measure the effectiveness of twitter opinion mining in terms of subjectivity, polarity and opinion target?

C. Identification of Relevant Literature

The strategy the researcher has used to construct the search strings was as follows:

1) The main terms are derived from the review questions based on the population,

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

intervention outcome and con- text.

2) Index terms mentioned in the articles are listed.
3) Search for alternative words.
4) Join synonyms and terms using Boolean OR.
5) Combine main terms from population, intervention and outcome using Boolean AND.

The complete search string initially used for the searching of the literature was as follows: ((Arabic OR Language- Independent OR Multilingual OR Bilingual OR Cross-lingual OR Language Independent OR Cross lingual OR multi-lingual OR corpus) And (Opinion mining OR Sentiment analysis OR emotion mining OR subjectivity analysis OR Lexicon-Based OR polarity of opinion OR opinion target identification OR fuzzy aspects of twitter opinion mining OR measurement of opinion mining OR measurement of subjectivity analysis OR measurement of polarity of opinion OR Experiment of opinion mining OR quality of opinion mining OR effectiveness of opinion mining) AND (technique OR Method OR process OR Approach) AND (feature OR Characteristic OR attribute OR aspect OR element) AND(Twitter OR microblog OR micro blog OR social media OR User-generated)). When using the complete search string defined above in the preliminary search, the researcher retrieved a very small number of articles. For instance, IEEEXplore retrieved only one article and CiteSeer five articles, respectively. Therefore, the researcher try simpler string than the one defined in the protocol to enable the retrieval of more results. The researcher used the keywords ”Opinion mining” OR ”Sentiment analysis” which resulted in a higher number of studies retrieved from various online databases. The primary search process involved the use of 7 online databases: ACM Digital library, EBSCOhost, IEEEXplore, ProQuest, Science-Direct, SpringerLink and Scopus. The selection of online databases was based on our knowledge of databases that index Opinion mining primary studies the researchers were aware of and the list of available online databases subscribed by Sudan University of Science and Technology (SUST). The researcher also searched the Citeseer website using similar keywords (i.e.,”Opinion mining” OR ”Sentiment analysis”); and online Google scholar was used to search for full text of articles. Upon completion of the primary search phase, the identification of relevant literature continued with the secondary search phase. During this search phase, all of the references in the papers identified from the primary sources were reviewed. If a paper was found to be suitable, it was added to the existing list of studies qualified for the synthesis.

D. Selection of Studies

The researcher inclusion criteria aimed to only include opinion mining and Sentiment analysis studies that targeted twitter. The literature search only covered studies published within the period of 2003-2013. The detailed inclusion criteria was composed of firstly studies that investigated features and techniques for twitter Sentiment analysis ,secondly studies that investigated features and techniques for Arabic Sentiment analysis of twitter and thirdly papers involving corpus on their subjects if the studies conducted are relevant to Arabic opinion mining. The main exclusion criterion consisted of papers not targeted micro blogging data. In addition, the following criteria were also applied:

1) Papers describing author(s) opinions without experi- ments or supporting evidence.
2) Papers describing opinion mining issues without high lighting techniques and features.
3) Papers describing tools (software or hardware) to sup- port opinion mining practice.
4) Papers describing opinions on specific field such as biomedical text mining
5) Papers describing opinions mining issues on document or reviews.
6) Papers not written in English and Arabic.

E. Data Extraction and Study Quality Assessment

To facilitate the data extraction process, a form was designed to gather evidence data relating to the research questions and measure the quality of the primary studies. When designing the studies' quality checklist, the researcher reused some of the questions proposed in the literature [10]. The checklist was composed of seven general questions (see Appendix B Table II) to measure the quality of both quantitative and qualitative studies according to the following ratio scale: Yes 1 point, No 0 points, and Partially 0.5 point. The

Journal of Theoretical and Applied Information Technology
20th October 2013. Vol. 56 No.2

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

resulting total quality score for each study ranged between 0 (very poor) and 7 (very good).

3. RESULTS

A. Introduction

In this section, the researcher presents the synthesis of evidence of the SLR, beginning with the analysis from the literature search results. During this phase the Science Direct database was chosen as the baseline database. Each article retrieved was compared with the existing list of papers in order to avoid duplication. The initial phase of the search process identified 19,677 studies using the ”twitter Opinion mining” OR ”twitter Sentiment analyses” search term as a total of all databases 401 out of them from Science Direct database. Of these, only 122 were potentially relevant based on the screening of titles and abstracts. Each of these studies was filtered according to the inclusion and exclusion criteria before being accepted for the synthesis of evidence. If titles and abstracts were not sufficient to identify the relevance of a paper, full articles were used. The researcher also carefully checked if there were any duplicate studies or if very similar studies were published in more than one paper. Based on the primary searches, 65 studies (53% of 122 studies) were ac- cepted for the synthesis of evidence after a detailed assessment of abstracts and full text and exclusion of duplicates. Most studies achieved above average quality scores and the others deemed very good and good quality. In the following section, the researcher presents the results for the SLR’s main research question and six subquestions. Each study is identified as Sn, where n represents the study’s number (see Appendix A for the list of studies used in this SLR).

B. Research Question

Question: ”What evidence is there of opinion mining studies conducted for twitter and Arabic tweets?” According to the result of the above question, the SLR identified 65 twitter opinion mining studies. The context of investigation varied via the sentiment analysis, opinion target identification and subjectivity identification for broadly type of targets. The SLR’s ultimate goal was to understand what techniques and features have been used in twitter sentiment analysis and is this technique and features used for Arabic tweets. (Of the 65 studies analyzed, two (3%) investigated opinion target identi- fication. Sixty one (94%) investigated polarity and subjectivity and two (3%) for Arabic corpuses.) In the following, the researcher is going to discuss the minor questions detail the SLR’s synthesis of evidence.

1) Subjectivity identification: ”What are the features and techniques that have been used for subjectivity identification of tweets? Are these features and techniques used for Arabic tweets?” Subjectivity identification means determine whether there is expresses opinions on text or not. Appendix B table III list the features used in subjectivity identification, studies that looked into each feature and whether the feature had used for Arabic tweets or not. Very few studies explicitly men- tioned subjectivity as a goal of their research. Three studies [S11], [S53], [S55] used combined features in subjectivity identification; these features were used also for Arabic tweets. However most of the studies handle the natural type of tweets as facts and emotional ones as subjectivity tweets then classify them into positive or negative ones and the rest ones use corpus of subjectivity words. The three studies used SVM as classification method.

2) Polarity identification: ”What are the features and tech- niques that have been used for polarity identification of tweets? Are these features and techniques used for Arabic tweets?” Polarity identification of the Subjectivity text refer to the Sentiment orientation classification of the text, which is done with assistance of polarity (Opinion) words. Appendix B table IV list the features used in polarity identification, studies that looked into each feature and whether the feature have been used for Arabic tweets or not. From table IV which deals with polarity studies by features the table shows that n- grams were the most common features investigated in polarity identification studies. Of the 61 sentiments analysis studies 33 (54%) used different types of n-grams features to classify tweets as positive, negative or natural. This features used in Arabic sentiments analysis for tweets by [S13], [S54]. The second most investigated features is a Combined features (e.g. POS n-grams, POS, named entity, emoticon and picture, etc.), 8 studies (13%) used those features, one study [S55] used Token (TOK), Lemma (LEM),Word forms, POS tagging, Standard Features (Unique, Gender, User ID) with a Polarity Lexicon for Subjectivity and Sentiment Analysis of Arabic Social Media. For TF-IDF and bag-of-unigrams there are 5 studies (8%)

Journal of Theoretical and Applied Information Technology
20th October 2013. Vol. 56 No.2

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

for each.3 studies (5%) used root word or tokens as features. 3 studies (5%) used a combination of n- grams and part of speech tags as features of their classifier. 2 studies (3%) used hash tags which is tweet specific feature. Only one study (2%) used lengthening phenomenon as a feature of their classifier. Findings showed a considerable improvement in accuracy. Data sparsity was the major prob- lems highlighted that affect classification accuracy. Based on classification method 12 studies (20%) used Support Vector Machines (SVM) as a method of classification [S2], [S3], [S6], [S11], [S33], [S12], [S19], [S44], [S55], [S16], [S52], [S53]. This method used for Arabic tweets by [S55]. 8 studies (13%) used combination methods for classification such as maximum entropy and k-nearest by [S14] for Arabic tweets; Naive Bayes (NB), and Support Vector Machines (SVM) by [S13] for Arabic tweets. Naive Bayes, Maximum Entropy, and Support Vector Machines by [S17], [S49].7 studies (11%) used Naive Bayes method [S32], [S40], [S45], [S46], [S48], [S42], [S31]. 5 studies (8%) used Corpus of tweets [S1], [S9], [S22], [S23], [S27]. 5 studies (8%) used lexicon-based classification method [S3], [s6], [S10], [S15], [S26]. Bagged decision trees for dictionary look-up is used by [S54] for Arabic tweets. See Appendix B table V for a list of studies by methods.

3) Target identification: What are the features and tech- niques that have been used for opinion target identification of tweets? Are these features and techniques used for Arabic tweets? Target identification of the Subjectivity text refers to entities and their attributes on which opinions have been expressed. Appendix B table VI list the features and methods used in target identification, studies that looked into each features and its method and whether a feature/method have been used for Arabic tweets or not. From table VI which deals with twitter opinion target studies by features/methods the table shows that only two studies [S7], [S9] deals with opinion target. [S7] used entity names and attributes as features to identify opinion target with UIMA 1 as method. [S9] used TF.IDF as a feature to identify opinion target with sentiment lexicon as method. No studies were found for opinion target identification of Arabic tweets.

4) Fuzzy aspects identification: What are the fuzzy aspects that have been studied for twitter opinion mining? Are they used for Arabic tweets? No studies were found for fuzzy aspects identification in twitter and Arabic tweets.

5) Arabic corpus identification: What are the Arabic corpus that have been used in Arabic opinion mining? Are they used for twitter? From table VII which deals with Arabic corpus for opinion mining the table shows that only two studies [S64], [S65] were founds and were not used for Arabic twitter sentiment analysis. Most twitter Arabic opinion mining studies made their own corpus for the purpose of the study [S55] then published it or not.

6) Metrics of effectiveness : What are the metrics that have been used to measure the effectiveness of twitter opinion mining in terms of subjectivity, polarity and opinion target? Referring to Appendix B table VIII effectiveness of twitter opinion mining approaches is measured by the accuracy, precision and recalls.The accuracy of approaches vary from good accuracy to poor accuracy depending on the features and methods have been used. 55 studies (87%) used accuracy,
7 studies (11%) used a combination of precision and recalls and only one study (2%) used precision.

4. DISCUSSIONS

For sentiment analysis, the most used classification method is the Support Vector Machines (SVM) for both Subjectivity and Polarity identification. The SVM has also been used for classification of Arabic tweets. The most used feature for the purpose is the n-grams, which has also been used for Arabic tweets. We also observe that only a little work has been done for target identification and development of Arabic corpus. More research and development need to be done on sentiment analysis system for Arabic tweets which can use a combination of features that are language-independent, Arabic-specific, domain-specific and twitter specific to obtain higher performance for opinion mining from Arabic tweets.

5. CONCLUSIONS

This paper described a systematic literature review (SLR) that investigated studies of twitter opinion mining and Arabic tweets. 65 primary studies were used, from which 61 studies focus on sentiment analysis, two studies focus on opinion target and two studies focus on Arabic

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

corpus. The SLR identified the current state of research and the features and techniques used for subjectivity identification, polarity identi- fication and opinion target identification. It can be concluded that features used for tweets opinion mining include n-gram, hash tags, lengthening phenomenon, TF-IDF, bag-of-word, part of speech and combined features and techniques to classify tweets include lexicon-based, corpus based, ontol- ogy learning, transfer learning, relaxation labeling, dictionary learning, probabilistic language models and classifiers based on one or combined methods that are SVM, NB, MaxEnt and k-nearest. For Arabic tweets, features used for include n-gram, Lemma, POS tagging, TF-IDF, standard features and root word and techniques to classify tweets include lexicon based method and machine learning methods that are bagged decision trees, polarity lexicon, classifier using SVM or MaxEnt and k-nearest method. More techniques and features can be explored to enhance the accuracy, precision and recall of Arabic tweets.

REFERENCES

[1] B. Liu, ”Sentiment Analysis and Subjectivity,” To appear in Handbook of Natural Language Processing, Second Edition, 2010

[2] Dan, Shen, C. Baudin, ”Sentiment Analysis in Practice,” ICDM11, December 12, 2011

[3] Saggion, Horacio, and Adam Funk. ”Extracting opinions and facts for business intelligence.” RNTI Journal, E (17) 119 (2009): 146.

[4] Jindal, Nitin, and Bing Liu. ”Mining comparative sentences and relations,” Proceedings of the National Conference on Artificial Intelligence. Vol. 5. No. 2. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 2006.

[5] (2013,Jun 8). The fastest, simplest way to stay close to everything you care about [Online]. Available:https://twitter.com/about

[6] Lunden, Ingrid. ”Analyst: Twitter Passed 500M Users in June 2012, 140M of them in US.” Jakarta BiggestTweetingCity.[Online] http://techcrunch.com/2012/07/30/analysttwitter-passed-500m-users-in- june-2012-140m-of-them-in-us-jakarta-biggest-tweeting-city (2012).

[7] de Haaff, Michelle. ”Sentiment Analysis, Hard But Worth It!, Customer- Think,” 2010.

[8] Flammia, Madelyn, and Carol Saunders. ”Language as power on the Internet.” Journal of the American society for information science and technology 58.12, 2007, pp.1899-1903.

[9] Versteegh, Kees, and C. H. M. Versteegh, The Arabic Language Columbia University Press, 1997.

[10] Habash, Nizar Y. ”Introduction to Arabic natural language processing,” Synthesis Lectures on Human Language Technologies 3.1, 2010, pp. 1-187.

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

APPENDIX A:

INCLUDED STUDIES REFERENCES (SN)

[1] L. Chen, W. Wang, and M. Nagarajan, ”Extracting diverse sentiment expressions with target-dependent polarity from twitter,” Proceedings of, 2012, pp. 5057.

[2] A. Bermingham and A. Smeaton, ”Classifying Sentiment in Microblogs: Is Brevity an Advantage? Categories and Subject Descriptors,” pp. 18331836, 1833.

[3] G. Paltoglou and M. Thelwall, ”Twitter, MySpace, Digg: Unsupervised Sentiment Analysis in Social Media,” ACM Transactions on Intelligent Systems and Technology, vol. 3, no. 4, Sep. 2012, pp. 119.

[4] P. H. C. Guerra et al, ”From Bias to Opinion: A Transfer-Learning Approach to Real-Time Sentiment Analysis,” 2011, pp. 150158.

[5] J. C. B. Rabelo, R. C. B. Prudłncio, and F. a. Barros, ”Leveraging relationships in social networks for sentiment analysis,” Proceedings of the 18th Brazilian symposium on Multimedia and the web - WebMedia 12, 2012, p. 181.

[6] V. N. Khuc et al, ”Towards building large-scale distributed systems for twitter sentiment analysis,” Proceedings of the 27th Annual ACM Symposium on Applied Computing - SAC 12, 2012, p. 459.

[7] C. Rodrguez-penagos, J. Grivolla, and J. C. Fib, ”A hybrid framework for scalable Opinion Mining in Social Media: detecting polarities and attitude targets,”, 2012, pp. 4652.

[8] D. Ramage, S. Dumais, D. Liebling, ”Characterizing Microblogs with Topic Models,” Proceedings of the Fourth International AAAI Confer- ence on Weblogs and Social Media, 2010.

[9] D. Spina, E. Meij, and A. Oghina, ”A Corpus for Entity Profiling in Microblog Posts,” LREC Workshop on , 2012.

[10] S. Brody and N. Diakopoulos, ”Cooooooooooooooollllllllll- lll!!!!!!!!!!!!!! Using Word Lengthening to Detect Sentiment in Microblogs,” 2011, pp. 562570.

[11] R. Passonneau et al, ”Sentiment Analysis of Twitter Data,” Proceedings of the , no. June 2011, pp. 3038.

[12] Y. Mejova, ”Sentiment analysis within and across social media streams,” 2012.

[13] A. Shoukry and A. Rafea, ”Sentence-level Arabic sentiment analysis,” 2012 International Conference on Collaboration Technologies and Sys- tems (CTS), May 2012, pp. 546550.

[14] A. El-halees, ”Arabic opinion mining using combined classification approach,” Proceedings of the International Arab Conference on Infor- mation Technology, ACIT. 2011

[15] M. Speriosu et al, ”Twitter Polarity Classification with Label propagation over Lexical Links and the Follower Graph,” Proceedings of EMNLP 2011, Conference on Empirical Methods in Natural Language Process- ing, Edinburgh, Scotland, UK, July 2731, 2011, pp. 5363.

[16] L. Barbosa and J. Feng, ”Robust Sentiment Detection on Twitter from Biased and Noisy Data,” Coling 2010: Poster Volume,Beijing, August 2010, pp. 3644.

[17] A. Go ,L. Huang,R. Bhayani, ”Twitter Sentiment Analysis,” June 6, 2009 [18] D. Davidov, O. Tsur, A. Rappoport ,”Enhanced Sentiment LearningUsing Twitter Hashtags and Smileys,” no. August, 2010, pp. 241249.

[19] D. Croce and R. Basili, ”Grammatical Feature Engineering for fine- grained IR tasks,” Proceedings of the Third Italian Information Retrieval Workshop IIR 2012, January 26-27, 2012, pp. 133143.

[20] A. Taheri et al, ”If You are Happy and Know It... Tweet,” wwws.cs.umn.edu.

[21] C. Mohan Dasari, D. Das, S. Bandyopadhyay, ”Topic Identification from Blog Documents: Roles of Bigram, Named Entity and Sentiment,” Language Engineering for Online Reputation Management, 26 May 2012

[22] A. Montejo-Rez et al, ”Ranked WordNet graph for Sentiment Polarity Classification in Twitter,” Computer Speech & Language, Apr.2013, pp. 115.

[23] J. Bollen and A. Pepe, ”Modeling Public Mood and Emotion: Twitter Sentiment and Socio-Economic Phenomena,” 2011, pp. 450453.

[24] G. Harihara, E. Yang, and N. Chambers, ”USNA: A Dual-Classifier Approach to Contextual Sentiment Analysis,” training.

[25] E. Kouloumpis, T. Wilson, and J. Moore, ”Twitter Sentiment Analysis: The Good the Bad and the OMG!,” 2011, pp. 538541.

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

[26] A. Minocha and N. Singh, ”Generating domain specific sentiment lexicons using the Web Directory,” vol. 3, no. 5, 2012, pp. 4551.

[27] Y. He, ”A Bayesian modeling approach to multi-dimensional sentiment distributions prediction,” Proceedings of the First International Work- shop on Issues of Sentiment Discovery and Opinion Mining – WISDOM 12, 2012,pp. 18.

[28] A. Hannak, E. Anderson, A. Mislove, and L. F. Barrett, ”Tweetin in the Rain: Exploring Societal-scale Effects of Weather on Mood,” 2010, pp.479482.

[29] K. Liu, W. Li, and M. Guo, ”Emoticon Smoothed Language Models for Twitter Sentiment Analysis,” 2010, pp. 16781684.

[30] E. Kontopoulos, C. Berberidis, T. Dergiades, and N. Bassiliades, ”Ontology-based sentiment analysis of twitter posts,” Expert Systems with Applications, vol. 40, no. 10, Aug. 2013, pp. 40654074.

[31] A. Bakliwal, ”Fine-Grained Opinion Mining from Different Genre of Social Media Content,” M. thesis. April, 2013.

[32] N. Bora, ”Summarizing Public Opinions in Tweets,” Journal Proceedings of CICLing, 2012.

[33] S. Mukherjee and P. Bhattacharyya, ”Sentiment Analysis in Twitter with Lightweight Discourse Analysis,” no. December 2012, pp. 18471864.

[34] A. Cui, M. Zhang, Y. Liu, and S. Ma, ”Emotion Tokens: Bridg- ing the Gap among Multilingual Twitter Sentiment Analysis,” no.20090002120005, 2011, pp. 238249.

[35] L. Zhang, R. Ghosh, M. Dekhil, M. Hsu, and B. Liu, ”Sentiment Analysis Combining Lexicon-based and Learning-based Methods for Twitter Sentiment Analysis,” 2011.

[36] J. Kim, J. Yoo, H. Lim, H. Qiu, Z. Kozareva, and A. Galstyan, ”Sentiment Prediction using Collaborative Filtering,” 2010.

[37] C. Zhang and J. Wang, ”Affective Computing Model for the Set Pair Users on Twitter,” vol. 10, no. 1, 2013, pp. 347353.

[38] V. Khuc, ”Approaches to Automatically Constructing Polarity Lexicons for Sentiment Analysis on Social Networks,” Thesis,2012.

[39] A. Davies, Z. Ghahramani, ”Language-independent Bayesian sentiment mining of Twitter TWITTER DATA , GEO MODELLING AND,” vol. 11, 2011.

[40] N. N. Bora and S. Bhattacharya, ”FEATURE BASED SENTIMENT Department of Computer Science & Engineering,” 2011.

[41] H. Saif, Y. He, and H. Alani, ”Alleviating data sparsity for twitter sentiment analysis,” The 2nd Workshop on Making Sense of, 2012, pp. 29.

[42] S. Narr, M. Hlfenhaus, and S. Albayrak, ”Language-Independent Twitter Sentiment Analysis,” dai-labor.de.

[43] S. Nepal, ”Linguistic Approach to Information Extraction and Sentiment Analysis on Twitter,” 2012.

[44] L. Jiang, M. Yu, M. Zhou, X. Liu, and T. Zhao, ”Target-dependent Twitter Sentiment Classification,” 2011, pp. 151160.

[45] H. Wang, D. Can, A. Kazemzadeh, F. Bar, and S. Narayanan,”A System for Real-time Twitter Sentiment Analysis of 2012 U . S . Presidential Election Cycle,” no. July 2012, pp. 115120.

[46] H. Saif, Y. He, and H. Alani, Semantic Smoothing for Twitter Sentiment Analysis, of the 10th International Semantic , 2011.

[47] R. de Groot , ”Data Mining for Tweet Sentiment Classification,” M. Thesis, 2012.

[48] T. Lake, ”Twitter Sentiment analysis,” 2011.

[49] A. Go, R. Bhayani, and L. Huang, ”Twitter sentiment classification using distant supervision,” CS224N Project Report, Stanford, 2009.

[50] C. Johnson, P. Shukla, and S. Shukla, ”On Classifying the Political Sentiment of Tweets,” cs.utexas.edu.

[51] A. Montejo-Raez, E. Martnez-Camara, M. T. Martn-Valdivia, L. A. Ure na-Lopez, ”Random Walk Weighting over SentiWordNet for Sentiment Polarity Detection on Twitter,” Proceedings of the 3rd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis,Jeju, Republic of Korea, 12 July 2012, pp. 310.

[52] A. Bakliwal, P. Arora, S. Madhappan,N. Kapre, M. Singh,V. Varma, ”Mining Sentiments from Tweets,” Proceedings of the 3rd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis, Jeju, Republic of Korea, 12 July 2012, pp. 1118.

[53] A. Agarwa and J. S. Sabharwal, ”End-to-End Sentiment Analysis of Twitter Data,” Proceedings of the Workshop on Information Extraction and Entity Analytics on Social Media Data,COLING 2012, Mumbai, December 2012, pp. 2934.

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

[54] L. Albraheem and H. S. Al-Khalifa, ”Exploring the problems of senti- ment analysis in informal Arabic,” Proceedings of the 14th International Conference on Information Integration and Web-based Applications and Services - IIWAS 12, 2012,p. 415.

[55] M. Abdul-mageed, K. Sandra, S. Kbler, and M. Diab, ”SAMAR: A System for Subjectivity and Sentiment Analysis of Arabic Social Media,” to Subjectivity and Sentiment , no. July 2012, pp. 1928.

[56] L. D. V. Tns, ”Do social media affect public discourses: A sentiment analysis of political tweets during the French Presidential Election campaign,” ,2012,pp. 121.

[57] V. M. K. Peddinti and P. Chintalapoodi, ”Domain Adaptation in Sentiment Analysis of Twitter,” Analyzing Microtext the 2011 AAAI Workshop (WS-11-05), 2011,pp. 4449.

[58] A. Pak and P. Paroubek, ”Twitter as a corpus for sentiment analysis and opinion mining,” Proceedings of LREC, 2010, pp. 13201326.

[59] H. Saif, Y. He, and H. Alani, ”Semantic sentiment analysis of twitter,” The Semantic WebISWC 2012, 2012.

[60] S. Wakade, C. Shekar, K. J. Liszka, and C. Chan, ”Text Mining for Sentiment Analysis of Twitter Data The University of Akron Department of Computer Science 2 . Previous Work.

[61] G. Gebremeskel, Sentiment Analysis of Twitter posts about news,” M thesis, University of Malta, May, 2011.

[62] D. Spina, E. Meij, M. de Rijke, A. Oghina, M. T. Bui, and M. Breuss, ”Identifying entity aspects in microblog posts,” Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval - SIGIR 12, 2012, p. 1089.

[63] A. Bifet and E. Frank, ”Sentiment knowledge discovery in twitter streaming data,” Discovery Science, 2010.

[64] M. Abdul-Mageed and M. Diab, ”AWATIF: A multi-genre corpus for Modern Standard Arabic subjectivity and sentiment analysis,” Proceed- ings of LREC, Istanbul, Turkey, 2012, pp. 39073914.

[65] M. Rushdi-saleh, M. T. Martn-valdivia, L. A. Urea-lpez, and J. M. Perea-ortega, ”OCA: Opinion Corpus for Arabic,” vol. 62, no. July2011, pp.20452054.

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

APPENDIX B

TABLE II: st u dy qua l i t y ch e c k l i s t

No	Item	Answer
1	Is it about opinion mining on twitter?	Yes/No
2	Are the aim(s) of study clearly stated?	Yes/No /Partially
3	Does the article have a stated research question or problem? Is it a valid research question in opinion mining of twitter?	Yes/No /Partially
4	Does the article describe the collection of new data? Is the data sufficient to answer the research question?	Yes/No /Partially
5	Does the article describe the procedure to follow in the study? Are the technique and features explained?	Yes/No /Partially
6	Were the approach to and formulation of the analysis was well conveyed?	Yes/No /Partially
7	Are the findings credible?	Yes/No /Partially

TABLE III: li s t o f f e at u r e s u s e d i n s u b j e c t i v i t y i d e n t i fi c at i o n s t u d i e s

No	Features	Studies	Arabic(yes/no)
1	Combined features(Token, Lemma, POS tagging, Standard Features)	S11, S53, S55	yes

TABLE IV: p o l a r i t y s t u d i e s b y f e at u r e s

No	Features	Studies	Arabic(yes/no)
1	n-grams	S1, S2, S3, S6, S11, S12, S24, S31,S29, S39, S18, S13, S15 S17, S43, S47, S38, S42, S50, S26, S62, S58, S56, S57, S60, S63, S61, S35, S27, S54, S46, S45, S40	Yes S13, S54
2	hashtags	S5, S52, S8	NO
3	lengthening phenomenon	S10	No
4	Combined features	S19, S44, S55, S59, S16, S21, S41, S37	Yes S55
5	TF-IDF weight of lemma	S22, S23, S51, S14, S4	Yes S14
6	word: (unique tokens or root or aspect/ feature)	S28, S30, S36	No
7	bag-of-unigrams(words)	S32, S48,S33, S34, S20	No
8	n-grams with part of speech tags.	S25, S49, S53	No

TABLE V: p o l a r i t y s t u d i e s b y m e t h o d

No	Method	Studies	Arabic(yes/no)
1	Corpus of unlabeled tweets, lemmas	S1, S8, S9, S22, S23, S27	No
2	Formal Concept Analysis, and (b)Ontology Learning.	S30	No
3	Lexicon-Based Classification	S3, S6, S10, S15, S26	No
4	Transfer learning approach	S4	No
5	collective classification Relaxation Labeling technique	S5	No
6	Naive Bayes (NB), and Support Vector Machines (SVM)	S13, S43, S47	Yes S13
7	k-nearest neighbors(kNN)	S18	No

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

8	Support Vector Machines (SVM) classifier	S2, S3, S6, S11, S33, S12, S19, S44, S55, S16, S52, S53.	Yes S55
9	Dictionary Learning(DL), Support Vector Machines (SVM), K Nearest Neighbors (KNN), Naive Bayes (NB)	S20	No
10	Named Entity Recognition (NER)SentiWordNet (SWN)	S21	No
11	Three way sentiment classifiers. adaboost.mh algorithm	S25	No
12	Dictionary based approach (Linux wamerican-small English dictionary)	S28	No
13	Polarity classification.	S29, S39	No
14	Naive Bayes	S32, S40, S45, S46, S48, S42, S31	No
15	Co-occurrence Graph.SentiWordNet	S34	No
16	Augmented lexicon-based method	S35	No
17	Linguistic Inquiry and Word Count. deep learning	S36	No
18	set pair analysis theory	S37	No
19	sentiment lexicons: graph based and Labeled TNG-based	S38	No
20	NB classifier from WEKA8 and the maximum entropy (MaxEnt) model from MALLET9	S41	No
21	Distant supervision: Naive Bayes, Maximum Entropy (MaxEnt), and Support-Vector Machines (SVM).	s17, S49	No
22	Maximum entropy classifier using label propagation	S50, S24	No
23	unsupervised approach :RandomWalk and SentiWordNet	S51	No
24	bagged decision trees. dictionary look-up	S54	Yes S54
25	maximum entropy and k-nearest	S14	Yes S14
26	Combined methods	S14, s17, S49, S41, S20, S13, S43, S47	Yes S14

TABLE VI: opinion target by features and methods

No	Features	Method	studie	Arabic(yes/no
1	Entity names, concepts or attributes asnoun, adjectival, adverbial or verbal phrases.	UIMA 1 (Ferrucci , Lally, 2004) architecture plus Solr-based clustering and indexing capabil ities.	S7	No
2	TF.IDF	Topic-specific sentiment lexicon.	S9	No