Developing resources for sentiment analysis of informal Arabic text in social media

ITANI, Maher, ROAST, Chris <http://orcid.org/0000-0002-6931-6252> and AL-KHAYATT, Samir

Published version

ITANI, Maher, ROAST, Chris and AL-KHAYATT, Samir (2017). Developing resources for sentiment analysis of informal Arabic text in social media. Procedia Computer Science, 117, 129-136.

Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 117 (2017) 129–136

Procedia Computer Science

www.elsevier.com/locate/procedia

3rd International Conference on Arabic Computational Linguistics, ACLing 2017, 5-6 November 2017, Dubai, United Arab Emirates

Developing Resources For Sentiment Analysis Of Informal Arabic Text In Social Media

Maher Itani, Chris Roast*, Samir Al-Khayatt

Communication and Computing Research Centre, Sheffield Hallam University, Sheffield, S1 2NU, United Kingdom

Abstract

Natural Language Processing (NLP) applications such as text categorization, machine translation, sentiment analysis, etc., need annotated corpora and lexicons to check quality and performance. This paper describes the development of resources for sentiment analysis specifically for Arabic text in social media. A distinctive feature of the corpora and lexicons developed are that they are determined from informal Arabic that does not conform to grammatical or spelling standards. We refer to Arabic social media content of this sort as Dialectal Arabic (DA) - informal Arabic originating from and potentially mixing a range of different individual dialects. The paper describes the process adopted for developing corpora and sentiment lexicons for sentiment analysis within different social media and their resulting characteristics. The addition to providing useful NLP data sets for Dialectal Arabic the work also contributes to understanding the approach to developing corpora and lexicons.

Peer-review under responsibility of the scientific committee of the 3rd International Conference on Arabic Computational Linguistics.

Keywords: sentiment analysis; corpora; lexicons; Arabic language; social media.

1. Introduction

Natural language processing applications work primarily with textual data, with objectives, such as, text categorization, machine translation and sentiment analysis. Their effectiveness is predicated upon the availability of a representative corpus for training, testing and validation. Such corpora need to embody information relevant to the intended language processing and also characterize - language as found "in the wild". Hence, corpora reflect the purpose for which they are used, for example they can be annotated with parts-of-speech (POS) tags, grammatical

* Corresponding author. Tel.: +44 -114-225-5555.
E-mail address: c.r.roast@shu.ac.uk

1877-0509
Peer-review under responsibility of the scientific committee of the 3rd International Conference on Arabic Computational Linguistics.
10.1016/j.procs.2017.10.101

elements (such as phrases, clauses and sentences). These can subsequently be used by various kinds of classifiers [1] such as Naïve Bayes (NB), decision tree (DT), Support Vector Machines (SVM), k-nearest neighbors (kNN), etc.. Classifiers are also designed and used for a variety of purposes, such as predicting movie sales, question answering, and other applications [2-10]. This also means that the corpora on which they rely on need to be domain specific. A corpora based on fashion reviews is likely to be inappropriate for classifying, say, movie reviews. For sentiment analysis classifiers aim to identify whether given posts are: positive, negative, neutral, etc.. Hence, enabling online product reviews to be assessed automatically.

In terms of language "in the wild", social media presents an interesting challenge since standard spellings and grammar are often ignored and, to some extent, new constructs are formed. This character of social media language also undermines the ease of annotating a corpus. It also limits the re-use and re-purposing of existing corpora - unless they are based on informal language use in the required domain.

This work focused upon the informal use of Arabic in social media as Dialectal Arabic (DA). Research related to building corpora is limited for the Arabic language when compared with English language. Authors in [9, 10, 11] attempt to partially fill this gap. However, Arabic resources become scarcer when we consider the sentiment classification of DA as that found in social media.

The paper describes building corpora and lexicons of social media in DA, using Facebook as a source. Section 2 provides a brief overview of the Arabic language and its characteristics, and currently available corpora. Section 3 describes forming annotated corpora and sentiment lexicons. Sections 4 and 5 review and discuss the outcomes.

2. The Arabic Language, Dialects and Corpora

The Arabic language is one of the top six major languages of the world [9, 10, 12]. The number of native speakers exceeds 200 million and it is the formal language in 22 countries. There are three different forms of Arabic [10]: Classical Arabic; Modern Standard Arabic (MSA) and Dialectal Arabic (DA). Classical Arabic is the language of Qur'an, the holy book in Islam, one of the world's major religions. MSA is the dialect used in education, books, television, newspapers, and in conversation among educated Arabs. DA is used to refer to the range of informal and local dialects. Such colloquial Arabic can be associated with geographical location, [13] provides the following rough groupings: Sudan and Egypt; Lebanon, Syria, Jordan, and Palestine; Gulf (Iraq, KSA, UAE, Kuwait, Qatar, Bahrain, and Yemen.); and, Libya, Tunisia, Algeria, and Morocco.

Since social media provides for open, weakly moderated expressions, local dialects are used. In addition, social media readily spans geographic and dialectical boundaries. Hence, our work treats DA as a dialect reflecting the diversity of different dialects found.

As a widely used language, Arabic has naturally been the focus of NLP. Existing research shows a number of different sources of Arabic corpora and techniques for deriving them. These include: mining data from databases [14-19], manual construction based on written text [20, 21], websites [22,23,24], as well as focusing upon social media posts [25,26], and spoken language [27]. Many Arabic corpora that may be used for text categorization, for example the publicly available Quranic corpus [28] that consists of one text file that includes syntactic and morphological annotation of the Quran. Al-Hayat Online Newspaper [29] and An-Nahar Online Newspaper [30] are two other Arabic corpora that cover different topics and whose texts are collected from online versions of the two newspapers.

Other corpora mentioned in sentiment analysis of Arabic text focus upon: web pages [38-39], documents [36], movie reviews [31] and social media [14,26,32-35,43, 45, 48].Authors in [46, 47] used different sources. So much so, that [31] propose translating existing tagged corpora found in other languages. Our work aims to help address lack of good corpora for sentiment analysis of informal Arabic with similar labels present in our corpus.

3. Corpus and Lexicon Development

The issues raised above motivated the development of our own corpora and sentiment lexicons for Arabic social media. The social media platform Facebook was chosen as a source because of its massive scale of usage - more than 1.4 billion users [40]. The objective of our sentiment analysis, is to support the classification posts as:

"negative", "positive", "dual", or "neutral". However, following a cursory assessment of social media posts, posts designed simply to drive web users lead to a fifth class of post, termed "spam".

To enable comparative assessment of our approach, two corpora were developed keeping to specific and distinct groups: news and arts (both capped at 1000 posts). The news corpus (NC) using posts collected from Al Arabiyya News Facebook page [41] and the arts corpus (AC) using posts collected from The Voice Facebook page [42]. The development of the corpora is described below. During the same process lexicons were developed for the two domains. In support of sentiment analysis lexicons were developed for the words and phrases (lexemes) within posts that could be interpreted as key to determining its sentiment. (Despite this, the sentiment of a lexeme may differ from the post in which it appears, see examples, below).

3.1. Data Collection

Corpora are either built using crawlers or collected manually. Although crawlers have the advantage of collecting large numbers of posts, they do require preprocessing to remove unwanted data [9]. In addition, social media platforms terms and conditions can constrain how their data is used. For example, in the case of Facebook, crawlers are not permitted. What is more, from a research ethics perspective, it is necessary to justify that those who posted to a group understood that their posts were in the public domain. In the case of our research this was supported through providing samples of the posts, their translation, and confirming the copyright conditions upon Facebook groups' posts. Subsequent use of the data collection adhered to good practice by anonymizing individual online identities and minimizing the risk of individuals posting being identifiable.

The posts consist of textual data posted by users as comments on posts written by the pages' administrators. The size of posts ranged from one word to a paragraph containing many sentences written in DA. Although DA includes different dialects, reflecting the non-localised nature of social media and its contributors, no attempt was made to differentiate dialects. Hence, the Facebook posts were treated as a reflection of the aggregate DA evident in social media. Out of interest, a later assessment of the posts indicated only 5% could be associated with a specific Arabic dialect.

3.2. Preprocessing

Posts were preprocessed based upon removing redundant content. Especially with short posts, social media users commonly repeat the same text more than once and duplicate others posts. Since the frequency of repeated posts does not influence their sentimental interpretation, repeated posts can be removed in preprocessing. In addition, posts were 'cleaned' by removing associated irrelevant data, such as time stamps and posted 'likes'. Both are best viewed as secondary data of no relevance to sentiment analysis. It is of interest to note that since the data was gathered Facebook has modified its "like" button to, what are termed, "Reactions" (see [44]). These include a range of predefined sentiment icons analogous to emoticons, yet, as with likes, they represent users' sentimental responses to textual posts and do not relate to the class of the post itself.

For the two domains following the preprocessing we found that: the 1000 arts based posts contained 12053 words (an average of 12 words per post), where the 1000 news based posts contained 8423 words (an average of 8 words per post)

3.3. Manual Tagging

Following pre-processing, expert native speakers tagged the collected posts. In our case four expert native Arabic speakers did the tagging. Their native dialect is Lebanese and they are experts in Egyptian, Syrian, and Palestenian dialects. Cases where a post's dialect was considered to be unfamiliar, passed to native speakers of the relevant dialect.

Fig. 1. Sample of a downloaded post in context.

The manual tagging employed the following rules:

- Posts expressing negative sentiments or feelings such as sadness, pessimism, hostility or any other negative feeling were classed as **negative**. For example:
  للاسف كان ذلك على حساب يسرى

  ("Unfortunately that was on Yusra's expense")

- Posts expressing positive sentiments or feelings such as enthusiasm, happiness, optimism, etc., were classed as **positive**. For example:
  مبروك مراد

  ("Congratulations Murad")

- If both positive and negative sentiments are expressed in the same post, posts were classed as **dual**. For example:
  مراد أخذ اللقب عن جدارة واستحقاق وموتوا بغيظكن ياحساد

  ("Murad deserves the title, die haters")

- If a post is inviting users to join or "Like" a Facebook page, it is classed as **spam**. For example:
  السلام عليكم ممكن تنشرون هذا البيج :

  https://www.facebook.com/pages/%D9%85%D8...

  ("Greetings, can you spread this page")

- If none of the above apply, a post is classed **neutral**.
  مراد شو شعورك ان ربحت احلى صوت وشو شعورك ان خسرت ؟

  ("Murad how would you feel if you win or lose the competition?")

Inter Annotator Agreement (IAA) was 97% which represents the percentage of posts classified similarly by all annotators. However, to strengthen the validity of the manual classification, only posts where all four annotators agreed were used. Hence, the IAA for the resulting corpora is 100%.

The resulting classes for the two domains (1000 posts in each) is shown in table 1. There is very little to differentiate the two domains, the greatest difference being in the dual and neutral classifications - that news (NC) has 6% more dual posts and the arts (AC) has 8% more neutral posts. This is slightly surprising since one might expect sentiment profile to differ between factual content and entertainment content.

Table 1. Frequency of posts of each class.

Paper	Arts (AC)	News (NC)	Total
Negative	224	230	454
Positive	233	236	469
Dual	151	161	312
Spam	197	193	390
Neutral	195	180	375

4. Lexicon Development

Conventionally, a lexicon provides a set of text strings (lexemes) for a specific domain that can be used to classify posts. A lexeme can be as small as one word or as long as sentence. The human taggers extracted the words and phrases from each post that they judged to be potentially influential upon its classification as positive, negative or spam. These were added as lexemes to a lexicon. Lexemes were extracted from the posts regardless of the posts actual classification. For instance, even if a post was classified as spam, it may still contain a positive phrase, the phrase would be recorded as a positive lexeme. For example, the following posts provide the positive lexemes مبروك ("congratulations") and جميل ("beauty") and

مبروك مراد
("Congratulations Murad")

عندك احلى صووووووووووووووت
("You have the most beautiful voice")

And, the following post contains the positive lexeme referring to beauty, despite being spam.
الآن انضم ,الفايسبوك على صفحة احلى
(The most attractive /lovely page on FB, join now)

Lexemes were then modified by factoring and extracting repeats, to improve recall:

- Factoring. This involves identifying new lexemes by removing suffixes and prefixes of existing lexemes. If the lexeme was جميلة (beautiful, used for females), the lexeme added to the set was جميل (beautiful, used for males). This also results in lexemes that are closer to word roots, and thus improves recall.
- Extracting repeats. What is more within social media and DA it was found that many lexemes involved regular informal departures from standard spelling. Consider the posts mentioned earlier, the word صووووووووووووووت (voice) is a wrong spelling variant of the word صوت, and clearly related variants include any number of repeated letters. Regular expressions have the ability to disregard repeated letters and thus were added as replacements for numerous individual lexemes. For example, if the lexeme جميل (beautiful) was converted to a regular expression, then this expression can detect all spelling variants such as جمممممميل , جميييييل and many more. (Hence, contrary to convention, “lexeme” in this research refers to the set of text phrases characterized by the regular expression.)

As far as we are aware this use of regular expression based patterns to capture repeats within lexemes is unique to this research, and provides a mechanism for helping to manage the informal character DA and social media content, when conducting sentiment analysis. Validating this approach the lexicons with regular expression patterns, and the lexicons without regular expression patterns where compared - applying them to an unseen corpus of 1000 posts. The lexicons with the patterns were found to have a 2% performance improvement.

With a view to validating the lexicons they were applied to the corpora and were shown to perform between 73% and 96%. The numbers of extracted lexemes show that social media posts constitute a good source to build an sentiment lexicon: out of the 2817 extracted lexemes, 2509 were unique yielding in ~11% redundancy rate. Although the corpus is relatively small, the numbers show that on average; at least one new lexeme can be extracted from each post. As for the upper threshold to this, a much bigger corpus needs to be annotated to see at which number of posts, or corpus size, will no new lexemes appear.

Table 2. Frequency of extracted lexemes.

	Arts	News	Total
Negative	743	678	1421
Positive	684	573	1257
Spam	96	43	139
Total	1523	1294	2817

A related question is whether the two lexicons reflect commonality within DA (independent of the domains). The lexicon commonality between domains is ~8% for Negative, ~14% for Positive and ~10% for spam. This suggests again either that the lexicons are far from comprehensive or that the two domains represent sentiment and opinions in very different ways. Further analysis using bigger corpora is needed to address these questions.

One other noteworthy domain difference is the lexicons' characterization of spam. Specifically, the spam frequency for news is less than half of that for the arts. Considering that the results are largely uniform, this appears significant. It suggests that spam is more easily characterized in the arts corpus than in the news corpus.

Other lexicons do exist for Arabic [48, 49], yet ours differ from them in two aspects: (1) source: our lexicon was constructed using data from social media, and (2) introduction of new label, the “spam”.

5. Lexicon and Corpora Characteristics

To assess the reliability of extracted patterns, we checked the extent to which lexicon characterizations matched those given in the corpora (see table 3). As we described earlier there are legitimate cases where characterization and classification do not match. However, the extent of the match indicates how easily DA can be classified based on simple patterns.

Table 3. Lexicon corpora consistency

	SA consistency with AC	SN consistency with NC
Negative	86%	90%
Positive	96%	96%
Spam	100%	100%

The 100% match for Spam shows the ease with which it can be identified. By contrast, the lower percentages for positive and negative can be attributed to language characteristics. Further similar analysis is to be conducted accommodating the classification inter-relations, such as the fact that 'Dual' was assessed on based on the presence of mixed sentiments (i.e. positive and negative).

The lexicon based identification of Spam was analyzed further, since spam posts may also include strong sentiments. Such cases were examined and it was found that spam patterns have greater dominance when compared other patterns in a post. This has implications for social media analysis of this sort and whether spam-like posts should be treated as categorically different.

The correlation between dialects and the class of lexicon or post was not examined because: the collection of posts did not consider the dialect, and therefore the dialects distribution is not balanced, and; the majority of posts used phrasing common to all dialects. Future work aims to conduct dialect specific analysis.

6. Conclusion

The Arabic language is a major world language, with informal forms of it used extensively on social media. This paper has provided an account of work developing corpora from Arabic Facebook posts [50]. While other corpora exist for text domain classification [14,26,32-35,43], support for sentimental analysis is limited. In addition to support sentiment analysis, the paper described how lexicons were constructed. The performance of the classifier was determined upon comparing its results against the annotation done by the human taggers. Details of implementation will appear in future publications. Although our work is focused on Facebook, the same process can be adopted when dealing with other social media, such as tweets (textual data posted on Twitter) and comments on Instagram, MySpace, Linkedin, etc.. Our analysis here, suggests that despite differing domains one might expect to find comparable results. A future research question is whether the resulting profile of DA on Facebook is similar in other platforms. For example, Twitter may differ due to its shorter post length and the resulting impact upon language used. As a simple example, Twitter users may be less inclined to emphasis through repetition such as can be seen in "صووووووووووووت". By contrast, LinkedIn appears to be based on longer professional posts closer to MSA. Clearly, any research based on adopting the approach described here will be subject to any social media's specific terms and conditions.

As a general account of corpora and lexicon construction for informal language, the phases of processing described here could be applied to other languages. Future research plans to examine whether similar processing is useful for sentiment analysis in other languages and whether analogous profiles are informative.

The other general finding of note is the distinctive nature of 'Spam' and what implications that has for sentiment analysis. One extreme would be treat the spam classification as a basis for excluding a post (as with, say, spam emails), whereas the alternative view would be accept that spam-like posts can express sentiments worthy of inclusion. Exploring this question in the future might require more careful consideration of how spam is classed and assessed along with the characteristics of social media use.

References

[1] El-Halees, A., 2011. Arabic opinion mining using combined classification approach.

[2] Jin, X., Li, Y., Mah, T. and Tong, J., 2007, August. Sensitive webpage classification for content advertising. In Proceedings of the 1st international workshop on Data mining and audience intelligence for advertising (28-33). ACM.

[3] Mishne, G. and Glance, N.S., 2006, March. Predicting Movie Sales from Blogger Sentiment. In AAAI spring symposium: computational approaches to analyzing weblogs (155-158).

[4] Shikalgar, N.R. and Badgujar, D., 2013. Online Review Mining for forecasting sales. International Journal for research in Engineering & Technologies (IJRET) December.

[5] Tatemura, J., 2000, January. Virtual reviewers for collaborative exploration of movie reviews. In Proceedings of the 5th international conference on Intelligent user interfaces (272-275). ACM.

[6] Somasundaran, S., Wilson, T., Wiebe, J. and Stoyanov, V., 2007, March. QA with Attitude: Exploiting Opinion Type Analysis for Improving Question Answering in On-line Discussions and the News. In ICWSM.

[7] Stoyanov, V., Cardie, C. and Wiebe, J., 2005, October. Multi-perspective question answering using the OpQA corpus. In Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing (923-930). Association for Computational Linguistics.

[8] Bollen, J., Mao, H. and Zeng, X., 2011. Twitter mood predicts the stock market. Journal of Computational Science, 2(1), 1-8.

[9] Izwaini, S., 2003, March. Building specialised corpora for translation studies. In Workshop on Multilingual Corpora: Linguistic Requirements and Technical Perspectives, Corpus Linguistics.

[10] The Arabic Language. 2013. [Online] Available at www.al-bab.com [Accessed 17 July 2016]

[12] Official Languages, Un.Org, United Nations, 2016. [Online] Available at: http://www.un.org/en/sections/about-un/official-languages/ [Accessed 17 July 2016]

[13] What is Spoken Arabic / the Arabic Dialects?, 2015, [Online] Available at: http://www.myeasyarabic.com/site/what_is_spoken_arabic.htm [Accessed 17 July 2016]

[14] Houngbo, H. and Mercer, R.E., 2014, June. An automated method to build a corpus of rhetorically-classified sentences in biomedical texts. In Proceedings of the First Workshop on Argumentation Mining (19-23).

[15] Lita, L.V., Schlaikjer, A.H., Hong, W. and Nyberg, E., 2005, July. Qualitative dimensions in question answering: Extending the definitional QA task. In PROCEEDINGS OF THE NATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE (Vol. 20, No. 4, 1616). Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999.

[16] Carlson, L., Marcu, D. and Okurowski, M.E., 2003. Building a discourse-tagged corpus in the framework of rhetorical structure theory. In Current and new directions in discourse and dialogue (85-112). Springer Netherlands.

[17] Samy, D., Sandoval, A.M., Guirao, J.M. and Alfonseca, E., 2006. Building a Parallel Multilingual Corpus (Arabic-Spanish-English). In Proceedings of the 5th Intl. Conf. on Language Resources and Evaluations, LREC.

[18] Dukes, K. and Habash, N., 2010, May. Morphological Annotation of Quranic Arabic. In LREC.

[19] Rytting, C.A., Rodrigues, P., Buckwalter, T., Novak, V., Bills, A., Silbert, N.H. and Madgavkar, M., 2014. ArCADE: An Arabic Corpus of Auditory Dictation Errors. ACL 2014, 109.

[20] Megyesi, B.B., Hein, A.S. and Johanson, E.C., 2006, May. Building a swedish-turkish parallel corpus. In Proceedings of the Fifth International Conference on Language Resources and Evaluation.

[21] El-Haj, M. and Koulali, R., 2013. KALIMAT a multipurpose Arabic Corpus. In Second Workshop on Arabic Corpus Linguistics (WACL-2) (22-25).

[22] Arabic Linguistic Blog. 2014. [Online]. Available at: http://archive.is/Ep1a [Accessed 27 Dec 2015]

[23] King Saud University Corpus of Classical Arabic. 2012. [Online] Available at: http://ksucorpus.ksu.edu.sa/?p=43. [Accessed 27 Dec 2015]

[24] Abdelali A, Cowie J, and Soliman H. 25th to 28th of July 2005, Building a modern standard Arabic corpus. In workshop on computational modeling of lexical acquisition. The split meeting. Croatia,

[25] Hamouda, S.B. and Akaichi, J., 2013. Social networks' text mining for sentiment classification: The case of Facebook'statuses updates in the 'Arabic Spring'era. International Journal Application or Innovation in Engineering and Management, 2(5), 470-478.

[26] Hamouda, A.E.D.A. and El-taher, F.E.Z., 2013. Sentiment analyzer for arabic comments system. Int. J. Adv. Comput. Sci. Appl, 4(3).

[27] Oostdijk, N., 1999. Building a corpus of spoken Dutch. In CLIN.

[28] Quranic Arabic Corpus, 2011. [Online] Available at: http://corpus.quran.com/download/default.jsp. [Accessed 7 Dec 2015]

[29] Al-Hayat Online Newspaper, 2011. [Online Available at: http://www.alhayat.com/. [Accessed 27 Dec 2015]

[30] Annahar Online Newspaper. 2015. [Online]. Available at: http://www.annahar.com/ [Accessed 27 Dec 2015]

[31] Farra, N., Challita, E., Assi, R.A. and Hajj, H., 2010, December. Sentence-level and document-level sentiment mining for arabic texts. In 2010 IEEE International Conference on Data Mining Workshops (1114-1119). IEEE.

[32] Refaee, E. and Rieser, V., 2014, May. An Arabic Twitter Corpus for Subjectivity and Sentiment Analysis. In LREC (2268-2273).

[33] Hajjem, M., Trabelsi, M. and Latiri, C., 2013. Building comparable corpora from social networks. In BUCC, 7th Workshop on Building and Using Comparable Corpora, LREC, Reykjavik, Iceland.

[34] Akra D. and Jarrar M., 2014, Towards Building a Corpus for Palestinian Dialect.

[35] Al-Sulaiti L and Atwell ES., 2006. The design of a corpus of contemporary Arabic. International Journal of Corpus Linguistics, 11(2), 135-171.

[36] Abdul-Mageed M, Diab MT, and Korayem M., 2011, Subjectivity and sentiment analysis of modern standard arabic. In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers-Volume 2, 587-591.

[37] Mustafa, M. and Suleman, H., 2011. Building a Multilingual and Mixed Arabic-English Corpus. In Proceedings Arabic Language Technology International Conference (ALTIC).

[38] Saad, M.K. and Ashour, W., 2010, November. Osac: Open source arabic corpora. In 6th ArchEng Int. Symposiums, EEECS (Vol. 10).

[39] Abdul-Mageed, M., Kübler, S., and Diab, M., 2012, 'SAMAR: a system for subjectivity and sentiment analysis of Arabic social media', in Proceedings of the 3rd Workshop in Computational Approaches to Subjectivity and Sentiment Analysis, 19–28.

[40] Arab Social Media Report, 2013, [Online]. Available at http://www.arabsocialmediareport.com/Facebook/LineChart.aspx?&PriMenuID=18&CatID=24&mnu=Cat [Accessed 27 Dec 2015]

[41] Al-Arabiya Facebook Page, 2011, [Online] Available at: http://www.facebook.com/AlArabiya. [Accessed 27 Dec 2015]

[42] MBCTheVoice Facebook Page, 2011. [Online] Available at: http://www.facebook.com/MBCTheVoice [Accessed 27 Dec 2015]

[43] Mubarak, H. and Darwish, K., 2014. Using Twitter to collect a multi-dialectal corpus of Arabic. ANLP 2014, p.1.

[44]Telegraph 2015 "Facebook to add happiness, love and anger buttons as alternatives to 'like'" [Online] Available at: http://www.telegraph.co.uk/technology/facebook/11918617/Facebook-to-add-emotional-reaction-buttons-as-alternatives-to-like.html [Accessed 7/1/2017]

[45] Nabil, M., Aly, M. A., & Atiya, A. F. (2015, September). ASTD: Arabic Sentiment Tweets Dataset. In EMNLP (pp. 2515-2519).

[46] ElSahar, H., & El-Beltagy, S. R. (2015, April). Building Large Arabic Multi-domain Resources for Sentiment Analysis. In CICLing (2) (pp. 23-34).

[47] Abdul-Mageed, M., & Diab, M. T. (2012, May). AWATIF: A Multi-Genre Corpus for Modern Standard Arabic Subjectivity and Sentiment Analysis. In LREC (pp. 3907-3914).

[48] Abdul-Mageed, M., & Diab, M. T. (2014). SANA: A Large Scale Multi-Genre, Multi-Dialect Lexicon for Arabic Subjectivity and Sentiment Analysis. In LREC(pp. 1162-1169).

[49] Badaro, G., Baly, R., Hajj, H., Habash, N., & El-Hajj, W. (2014). A large scale Arabic sentiment lexicon for Arabic opinion mining. ANLP 2014, 165.

[50] Itani, M. (2017). Corpus of Arabic social media posts manually classed for sentiment analysis. SHU Research Data Archive (SHURDA). http://doi.org/10.17032/shu-170008