Arabic SentiWordNet in Relation to SentiWordNet 3.0

Samah Alhazmi *alhazmis@cs.man.ac.uk*
Text Mining Research Group
School of Computer Science
University of Manchester
Manchester, UK, M13 9PL

William Black *william.black@manchester.ac.uk*
School of Computer Science
& National Centre for Text Mining
Manchester Institute of Biotechnology
University of Manchester
Manchester, UK, M1 7DN

John McNaught *John.McNaught@manchester.ac.uk*
School of Computer Science
& National Centre for Text Mining
Manchester Institute of Biotechnology
University of Manchester
Manchester, UK, M1 7DN

Abstract

Sentiment analysis and opinion mining are the tasks of identifying positive or negative opinions and emotions from pieces of text. The SentiWordNet (SWN*)* plays an important role in extracting opinions from texts. It is a publicly available sentiment measuring tool used in sentiment classification and opinion mining. We firstly discuss the development of the English SWN for versions 1.0 and 3.0. This is to provide the basis for developing an equivalent SWN for the Arabic language through a mapping to the latest version of the English SWN 3.0. We also discuss the construction of an annotated sentiment corpus for Arabic and its relationship to the Arabic SWN.

Keywords: Opinion Mining, Sentiment Analysis, WordNet, SentiWordNet, Arabic.

1. INTRODUCTION

Text mining involves the automated extraction of information from texts, often from large volumes of texts. An area of growing interest within text mining is opinion mining, which involves assessing whether a text is objective or subjective, and, if it is subjective, whether it is positive or negative [1, 2, 3, 4]. This is relevant to many tasks such as determining public opinion about a particular product, or tracking movements in public opinion in relation to questions of public policy. It involves both determining the polarity of a text (if any) and the strength of the text polarity. Texts may be weakly, mildly or strongly positive or negative and these differences can be highly relevant to the conclusions that can be drawn from the analysis.

Given the intrinsically subjective nature of opinions, assessing the quality of the results generated by any tool raises particular difficulties [1, 5].

This paper sheds light on the development of SentiWordNet (SWN) 1.0 and 3.0, a publicly available resource used in sentiment classification and opinion mining [1], and describes and compares the effectiveness of the English versions. SWN is an evolving resource that maps to consecutive versions of WordNet (WN) [15] — a resource consisting of synsets[1] that list (disambiguate) the multiple senses of words which are often ambiguous. This means that they may have multiple senses and vary in meaning by context. These disambiguated words are then glossed (explained), and it should be noted that the words in the gloss are not themselves disambiguated in the original WN.

SWN can also be applied across different languages but, as it maps to WN, it requires an appropriate version of WN for the language in question.

In the case of Arabic, this requires the development of an Arabic version of WN 3.0 that is then mapped to the English WN 3.0. In the absence of this, the development of sentiment analysis for Arabic texts will fall behind in one of the most promising areas for text mining. What is more, the existing Arabic WN 2.0 does not contain an extensive range of synsets and this inadequacy needs to be addressed rapidly. Based on this fact, this paper contributes toward presenting an Arabic SWN in relation to the latest version of the English SWN 3.0, taking into account upgrading the Arabic WN 2.0 to version 3.0.

Section 2 gives a brief summary of lexicons that are available for sentiment. Section 3 draws a brief history of SWN and discusses the English versions 1.0 and 3.0. Section 4 then outlines the preparatory steps that need to be taken to develop a version of SWN to support analysis of Arabic texts, and includes a discussion of the mapping process between the English and the Arabic versions. Section 5 discusses related work. Finally, future directions are briefly considered in the conclusion, section 6.

2. BACKGROUND

Numerous lexicons are available for sentiment analysis. There can be low or quite significant degrees of disagreement among them. It is possible to intentionally use such contrasts in solving conflicts or, alternatively, they may be accepted to be genuine areas of uncertainty.

WordNet may be utilised to determine useful lexicons from small seed sets, including in cases where the differences are not clearly encoded within WordNet.

One of the key benefits of lexical induction is the ability to include domain-specific effects. Table 1 summarises a number of lexicons and their characteristics.

Lexicons	Brief Descriptions
SentiWordNet [1, 5, 15]	SentiWordNet assigns positive or negative numerical sentiment values to WordNet synsets. It is available for free as long as it is not used for financial gain, with business users required to seek a license.
Liu's Opinion Lexicon [22, 23]	It provides for spelling errors, morphological variation, vernacular and internet terminology.
MPQA Subjectivity Lexicon [24]	The Multi-Perspective Question Answering (MPQA) Subjectivity Lexicon is distributed under the GNU Public License.
Harvard General Inquirer [25]	The Harvard General Inquirer is a lexicon which assigns syntactic, semantic as well as practical data to part-of-speech tagged words.
LIWC [26]	Linguistic Inquiry and Word Counts (LIWC) is a proprietary database of classified common terms. It is priced at $90 and has categories which are very close to the ones in the Harvard General Inquirer.

TABLE 1: Summary of number of lexicons and their characteristics.

[1] Synsets are sets of "senses" for a word that clarify the particular sense in which it can be used. The words that clarify each sense are called the "gloss".

The aforementioned lexicons have only basic divergence in categorisation. They possess word banks that are not the same and, thus, can only be contrasted to a limited extent.

For the purposes of this paper, we have chosen SentiWordNet as a sentiment lexicon for our work, to build an Arabic version of this lexicon, as will be discussed later.

3. SENTIWORDNET: BRIEF HISTORY

Researchers have attempted to develop systems to automatically label words that indicate opinions as being either positive or negative [6,7,8]. The prior and related question of whether a word is in fact a marker of opinion or not, whether it is subjective or objective, has received less attention [9]. Early attempts involved labelling words without making distinctions between the different senses in which a word may be used, with the result that the word rather than its sense is classified. This has limitations as the same word very often has multiple senses and any system that fails to capture these variations in meaning is severely restricted in functionality and reliability.

2.1 SentiWordNet 1.0

Esuli and Sebastiani [1] attempted to address this limitation by developing a resource, SWN (version 1.0), that goes beyond simply listing and classifying a word to classify the sense in which it is being used. That is, one word may — and usually does — have multiple senses. In the WN world, such multiple senses for a word are called synsets.

In SWN 1.0, Esuli and Sebastiani augmented each WN synset by assigning a numerical score in three categories: Obj (Objective), Pos (Positive) and Neg (Negative). These mean that a synset is assessed as being either objective or subjective, and, within the subjective category either positive or negative. As this scheme is applied to synsets, it addresses the issue that the same word may have different senses that have different opinion-related properties.

The sum of the three scores is always 1.0. This allows the score to reflect the fact that a synset may have opinion-related properties to a certain degree. An example given by Esuli and Sebastiani is the synset [estimable (3)] with the sense “may be computed or estimated”. This has an Obj score of 1.0, and Pos and Neg scores of 0.0. This indicates that the sense is clearly classified as non-subjective. In contrast, the synset [estimable (1)] with the sense (gloss) “deserving of respect or high regard” has a Pos score of 0.75, a Neg score of 0.0 and an Obj score of 0.25.

In effect, this method allows senses to be classified as having multiple aspects to varying degrees. This idea was originally introduced by Kim and Hovy [8], but Esuli and Sebastiani’s SWN gives it functionality. They argue that the application of tools that grade opinion-related properties will play an important role in the future of text mining.

2.2 SentiWordNet 3.0

SWN 3.0 is a development of SWN 1.0 [5]. The intermediate development stages SWN 1.1 [11] and SWN 2.0 [12] were primarily for internal use of the developers and it is standard to compare SWN 3.0 directly with its public predecessor SWN 1.0. There are three primary differences between SWN versions 1.0 and 3.0:

a) Version 3 was developed as an annotation of WN 3.0 whereas version 1 applied to WN 2.0. This amounts to an updating of the resource, and has implications when it comes to comparing the accuracy of version 3.0 with version 1.0.

b) The main difference and the most important one for improving the accuracy of the resource is that an additional analytical stage is added after the semi-supervised learning stage associated with SWN 1.0. This is an iterative random-walk step that is carried out on the result of the semi-supervised learning algorithm.

c) In SWN 1.0, the glosses from WN 2.0 are used in the training stage, not the synsets themselves, with the result that the classifier is a gloss classifier rather than a synset classifier. In SWN 1.0, the gloss is a non-disambiguated collection of words. Esuli and Sebastiani call this a "bag of words" approach, where no attempt is made to determine the sense of the words (to disambiguate them). They are simply given a frequency-weighted score. This is different from SWN 3.0, where the development of the random-walk step requires the gloss to be disambiguated: in other words, to yield a further collection of synsets. Thus, in contrast to SWN 1.0, SWN 3.0 construction uses a "bag of synsets" approach.

A. Random-Walk step: The Concept

This step makes an assumption about relationships between words. The definiens (the words defining the word in question) is in a binary relationship with the definiendum (the word being defined). The assumption that underlies the random-walk step is that a direct link can be posited between synset 1 (S1) and synset 2 (S2), if and only if S1 occurs in the definiens of synset 2. The idea underlying this assumption is that if many words in the definiens are positive (or negative) then it is plausible that the definiendum is positive (or negative) too. This assumption allows positivity (or negativity) to, in the words of Baccianella et al. [5], "flow through the graph from the words used in the definitions to the words being defined."

B. Difficulty in implementing the random-walk step

A major issue is that, in WN, although the definiendum is a synset, the words used in the glosses are not. That is, the definiendum is disambiguated (different senses of the word are separated out), whereas the words in the definiens are non-disambiguated (the sense in which the word is being used is not made clear). This is not suitable for the random-walk step, which implies links between the words in both definiens and definiendum. For this to work, the glosses themselves must consist of a string of synsets (disambiguated words).
In the case of SWN 3.0, the Princeton WordNet Gloss Corpus[2] was used to achieve this. This contains manually disambiguated glosses for WN 3.0.

C. Assessment of SentiWordNet 3.0 Including the Random-Walk step, in relation to SentiWordNet 1.0

Baccianella et al. assessed whether the addition of the random-walk step improves or reduces the accuracy of the results delivered by SWN 3.0 in comparison with SWN 1.0, which lacked this element. Both SWNs were assessed using a small manually annotated subset of WN which was then compared with the automatic annotations of the same synsets by the respective versions of SWN.

The methodology used in their assessment is described below. There are two main difficulties that the methodology had to address:

a) How to create an "objective" criterion or "gold standard" against which the results of the SWN can be compared.

b) How to make an assessment that has validity across different versions of WN.

SWN 1.0 was evaluated on Micro-WN (Op) [16] which consists of 1,105 synsets of WN 2.0 that were manually annotated for sentiment by 5 people. The methodology involved the 5 annotators

[2] http://wordnet.princeton.edu/glosstag.shtml

working together to develop a common assessment understanding, and then working individually to increase speed. All synsets were rated by more than one annotator and results averaged. As with SWN 1.0 itself, they scored for Pos, Neg and Obj with a rating that added up to 1.0. The manual results can then be compared with SWN 1.0.

The difficulty with applying the same method to an assessment of SWN 3.0 is that SWN 3.0 applies to WN 3.0, which has different synsets from WN 2.0. To make a meaningful assessment of relative accuracy, the synsets that were manually assessed in WN 2.0 must first be mapped across to WN 3.0.

Baccianella et al. acknowledge the limitations of this process. They used three tools that were applied consecutively, with a later tool only being used where an earlier one failed to produce a result. The three tools are:

a) WN sense mappings (nouns and verbs only).

b) Synset word matching (if a synset contains the same words in Micro-WN(Op) and WN 3.0, and uniquely so, then they are considered to describe the same concept).

c) Gloss similarity (the greatest similarity between glosses determines the most likely equivalence of sense). They examined some results manually and found them to be satisfactory, but recognise that the results of the mapping process have not been completely checked for correctness: this would imply a complete search of WN 3.0 synsets of the same polarity to find the best match for each word in Micro-WN(Op).

D. Issues concerning SentiWordNet 3.0

The matching process allows the results of SWN 1.0 and SWN 3.0 to be compared. The rankings for Positivity and Negativity between SWM 1.0 and SWN 3.0 that were in complete agreement with the manual ranking would have a value of 0, complete disagreement a value of 1. Thus, the lower the value, the greater the agreement.

SWN 3.0 appears significantly more reliable than version 1.0 with a 19.48% increase in ranking by positivity and 21.96% by negativity. Even given the qualifications that were discussed above, this appears to be a marked improvement.

Important issues are raised by the development of SWN 3.0:

a) It is difficult to establish a gold standard against which to measure the effectiveness of an opinion mining resource. This becomes an even greater problem as the development of the target resource (WN) changes. Comparisons are difficult and likely to become more so.

b) SWN 3.0 has evolved by an additive process, adding steps. There is room for debate whether future development will be best achieved by continuing the same process — adding steps that refine the results further — or by revisiting and refining some existing steps.

4. GENERATING ARABIC SENTIWORDNET

Arabic is a widely used language that has both economic and political importance. It is natural that tools that enable sentiments in Arabic texts to be extracted and assessed are of great interest. Existing development work has been focused on developing for Arabic the equivalent of the resources that exist for English. By adding sentiment information to the Arabic WN to generate the Arabic SWN, we have improved the set of WN-based resources for the benefit of researchers in Arabic Natural Language Processing (NLP). This has the potential to enable future opinion mining tools to be developed directly for Arabic texts.

The key initial requirement for the application of SWN to Arabic is a version of WN on which it can operate. There is an existing Arabic version of WN 2.0 [13, 14] — Arabic WN 2.0[3], but not of WN 3.0.

Therefore, we built the database for Arabic SWN taking into account all the preparation levels shown in Figure 1:

a) The Arabic WN 2.0 database was upgraded to version 3.0 by mapping to the latest English WN 3.0 database.

b) The English SWN 3.0 database was also mapped to the new version of the Arabic WN 3.0 database.

c) All fields in our new database — Arabic WN 3.0 — were checked and revised with the English SWN 3.0, and then only the fields that existed in the English version (which express sentiments) were kept in the Arabic version to give us the Arabic SWN, with the rest deleted.

Through this mapping, an Arabic SWN database was generated. However, it contains fewer words and there is still a need to translate all the other fields that exist in the English version. Thus, the total number of words existing in the Arabic SWN is around 10,000, which includes verbs, adjectives, adverbs and nouns.

3.1 Approaches for using the Arabic SWN

Two different approaches can be used for developing the Arabic SWN:

a) Using the database we built as a multi-lingual setup to be applied to both English and Arabic contexts.

b) Having all synsets in the English version translated into the Arabic version – which is the approach we used in our work.

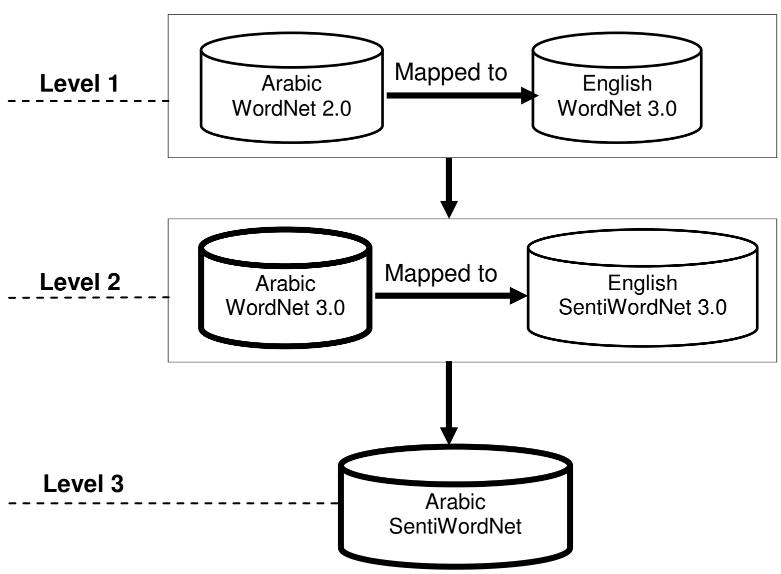

FIGURE 1: Preparation step for generating Arabic SWN.

[3] http://www.globalwordnet.org/AWN/

3.2 The AWN-WordNet Mappings

When the AWN was constructed, each of its synsets was mapped to Princeton WordNet (for English) Version 2.0 despite two more recent versions of Princeton WordNet being released while the AWN was under construction. However, SentiWordNet provides additional annotations on WordNet version 3.0, so it was necessary to update the Arabic to English mapping to WordNet 3.0. This was done using the mapping from WordNet 2.0 to 3.0 available from http://nlp.lsi.upc.edu, which was constructed automatically using the procedure described in [30]. The mappings give a unique WN3.0 synset ID for 99.7% of WN2.0 adverb synsets, 98.77% of adjective synsets, 99.39% of noun synsets and 98.92% of verb synsets. Of the ambiguously mapped WN2.0 synsets, a majority have no AWN linkage, but some remained to be verified manually.

3.3 Methodology

Several research studies have been done on opinion mining and sentiment analysis using reviews and movies as their datasets, either in English or Arabic [20, 27, 28, 29]. For the purpose of our research, to generate a new corpus for Arabic sentiment analysis, the data was generated from Arabic social media in the form of technology blogs (2,350 sentences). Technology blogs provide various challenges:

a) Use of a foreign language (English in our case) for names of technologies, companies, software or programs.

b) Transliterations within the texts.

c) Difficulty of recognizing opinions and sentiments expressed to show users' opinions towards these technologies and companies.

One of the purposes of building this corpus is to evaluate the sentiment coverage of the Arabic SWN. Our corpus construction methodology involves the manual detection of sentiments via manual annotation, which was carried out according to our annotation guidelines, which were:

a) Determine positive or negative words about companies.

b) Determine positive or negative words about technologies, products or systems.

c) For each sentence, each sentiment word should be detected as positive or negative and then determination is made whether the whole sentence is positive or negative.

Finally, the annotated sentiment words are compared with the words that exist in the Arabic SWN for the evaluation step (section 3.4). Tables 3 and 4 show some Arabic examples of the sentiment annotation tasks about companies and technologies, products or systems, respectively (note: all examples are translated into English for ease of understanding):

Companies	
A positive sentence	***A negative sentence***
خطوة كبيرة من الفيس بوك بتوفير هذا المتجر والذي سينجح بكل سهولة.	كاسبرسكي: أبل تفتقد للجدية في حماية نظام الماك
A big step from Facebook to provide this store, which will succeed easily.	Kaspersky: Apple is not serious about protecting the Mac system.

TABLE 3: Examples of the sentiment annotation tasks about companies.

Technologies, products or systems	
A positive sentence	***A negative sentence***
أحب استخدام خدمة فورسكوير لتحديد موقعي.	شخصياً أرى أن مستقبل نظام الاندرويد بدأ يصبح مخيفاً وغير آمناً.
I love using foursquare to determine my location.	Personally, I think the future of Android is becoming scary and unsecure.

TABLE 4: examples of the sentiment annotation tasks about technologies, products or systems

3.4 Discussion of Annotation Results

It was found that 31% of sentences were annotated with negative sentiment and 64% with positive. The inter-annotator agreement (IAA) is shown in figure 2.

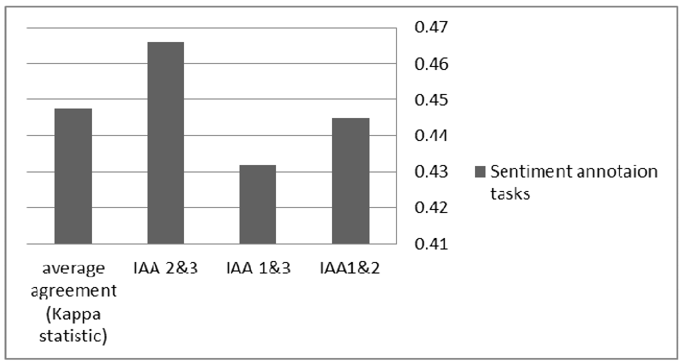

FIGURE 2: The Inter-annotator agreement (IAA) and average agreement (Kappa statistic)

Three annotators were involved and the Kappa statistic was used to calculate IAA for each sentence: the overall result obtained was 0.45 which is considered as "moderate agreement" [17].

We intend to improve on this result by adopting a cyclic annotation approach. However, we must first establish the sentiment coverage of the Arabic SWN. We therefore carried out an evaluation step to check which annotated sentiment words occur in the Arabic SWN. We already know that the Arabic SWN includes fewer synsets than English SWN. It was established that around 5% of annotated sentiment words did not occur in the Arabic SWN. This result is clearly influenced by the size and nature of our corpus. However, taken together with the difference in the number of synsets between Arabic SWN and the English SWN, it is evident that we must expand the coverage of Arabic SWN. Returning to our cyclic annotation approach we plan to speed up and improve annotation by incorporating automatic look up of Arabic SWN to tag recognised sentiment words for validation by the annotators, which will reduce the annotation burden. We will though firstly translate the missing synsets in Arabic SWN from English SWN. As this is a translation exercise it is to be expected that there will be Arabic specific sentiment synsets that will be still missing. Some of these may be added from our corpus, once we are able to determine the complete coverage of the Arabic SWN augmented by the translated synsets.

5. Related Work

Due to the lack of an existing SentiWordNet lexicon for the Arabic language, several research studies on opinion mining and sentiment analysis for Arabic have used lists of sentiment words to cover their research needs [18, 19, 20]. Recently, a study done by Abdul-Mageed and Diab [21] focussed on expanding a polarity lexicon of Modern Standard Arabic built manually by leveraging various existing polarity lexica for English. The utility of their expanded lexicon was not tested and moreover they focussed on adjectives.

Unfortunately, a comparison cannot be made between our Arabic SWN and previous systems because they have not yet been made available; furthermore, our database links to the original Arabic WordNet which is not the case for other previous systems.

6. CONCLUSION AND FUTURE DIRECTIONS

SWN has proved to be an effective tool for opinion mining and sentiment classification, and the improvement in effectiveness that has been achieved between versions 1.0 and 3.0 gives grounds for optimism that it provides a tool with further developmental possibilities. That is, it is flexible enough to accommodate both additional refining steps and new work on the existing steps.

In this paper, we reported efforts to generate an Arabic SWN database in relation to the English SWN 3.0. There is however a need to increase the number of words in our Arabic SWN. We plan also to refine further the annotation of our corpus to achieve a higher IAA score, especially as we plan to use this corpus for research purposes in Arabic sentiment analysis. The ranking process for positivity and negativity is another step to be taken in the near future. All these limitations above need to be resolved in future to enhance the final version of our Arabic SWN which will be made publicly available.

Acknowledgments

This research was supported via a PhD scholarship from Saudi Cultural Bureau.

7. REFERENCES

1. A. Esuli, and F. Sebastiani, “SENTIWORDNET: A Publicly Available Lexical Resource for Opinion Mining.” in Proc. of the 5th Conference on International Language Resources and Evaluation (LREC’06), 2006, pp. 417–422.

2. B. Pang, and L. Lee, “A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts.” in Proc. of the ACL, 2004, pp. 271–278.

3. B. Liu, “Opinion Mining and Sentiment Analysis.” in B. Liu, Web Data Mining, Springer Berlin, Heidelberg, 2011, pp. 459–526.

4. T. Wilson, J. Wiebe, and R. Hwa, “Just how mad are you? Finding strong and weak opinion clauses.” in Proc. of AAAI, 2004, pp. 761–769.

5. S. Baccianella, A. Esuli, and F. Sebastiani, “SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining.” in Proc. of the 7th Conference on International Language Resources and Evaluation (LREC’10), 2010, pp. 2200–2204.

6. H. Takamura, T. Inui, and M. Okumura, “Extracting Emotional Polarity of Words using Spin Model.” in Proc. of ACL-05, 43rd Annual Meeting of the Association for Computational Linguistics, 2005, pp. 133–140.

7. A. Esuli and F. Sebastiani, “Determining the semantic orientation of words through gloss classification.” in Proc. of the 14th ACM international conference on Information and Knowledge Management ACM, 2005, pp. 617–624.

8. S.-M. Kim and E. Hovy, “Determining the sentiment of opinions.” in Proc. of the 20th international conference on Computational Linguistics Association for Computational Linguistics, Geneva, Switzerland, 2004, pp. 1-1367.

9. A. Esuli, and F. Sebastiani, “Determining Term Subjectivity and Term Orientation for Opinion Mining.” in Proc. of the 11th Conference of the European Chapter of the Association for Computational Linguistics (EACL’06), 2006, 193-200.

10. A. Valitutti, C. Strapparava and O. Stock, “Developing Affective Lexical Resources.” Psychology journal, 2004, pp. 61–83.

11. A. Esuli and F. Sebastiani, “SentiWordNet: a high-coverage lexical resource for opinion mining.” Institute of Information Science and Technologies (ISTI) of the Italian National Research Council (CNR), 2006.

12. A. Esuli, “Automatic generation of lexical resources for opinion mining: models, algorithms and applications.” SIGIR Forum, ACM, vol. 42(2), 2008, pp. 105–106.

13. W. Black, S. Elkateb, and P. Vossen, “Introducing the Arabic WordNet Project.” in Proc. of the third International WordNet Conference (GWC-06), 2006, 295-300.

14. S. Elkateb, W. Black, D. Farwell, P. Vossen, A. Pease, and C. Fellbaum, “Arabic WordNet and the Challenges of Arabic.” In Proc. of the British Computer Society Conference on Challenges of Arabic for Natural Language Processing and Machine Translation.The British Computer Society (BCS), London, 2006, pp. 15–25.

15. G. A. Miller, R. Beckwith, C. Fellbaum, D. Gross and K. Miller. "Introduction to wordnet: An on-line lexical database*." International journal of lexicography, vol. 3(4), 1990, pp. 235-244.

16. S. Cerini, V. Compagnoni, A. Demontis, M. Formentelli, and G. Gandini. "Micro-WNOp: A gold standard for the evaluation of automatically compiled lexical resources for opinion mining," Language resources and linguistic theory: Typology, second language acquisition, English linguistics, A. Sans`o, Ed. Milano: IT, 2007, pp. 200–210 .English linguistics, pp. 200–210. Franco Angeli Editore, Milano, IT.

17. A. J. Viera and J. M. Garrett. "Understanding interobserver agreement: the kappa statistic." Family Medicine, vol. 37, 2005, pp. 360–363.

18. A. A. Al-Subaihin, H. S. Al-Khalifa and Al-Salman, A. S. "A proposed sentiment analysis tool for modern Arabic using human-based computing." In Proc. of the 13th International Conference on Information Integration and Web-based Applications and Services, Ho Chi Minh City, Vietnam, ACM**,** 2011, pp. 543-546.

19. N. Farra, E. Challita, R. Abou Assi and H. Hajj. "Sentence-Level and Document-Level Sentiment Mining for Arabic Texts." in Proc. of the 2010 IEEE International Conference on Data Mining Workshops, IEEE Computer Society, 2010, pp. 1114-1119.

20. M. Rushdi-Saleh, M. T. Martín-Valdivia, L. A. Ureña-López and J. M. Perea-Ortega. "OCA: Opinion corpus for Arabic." J. Am. Soc. Inf. Sci. Technol, vol. 62, 2011, pp. 2045-2054.

21. M. Abdul-Mageed and M. Diab. "Toward Building a Large-Scale Arabic Sentiment Lexicon." in Proc. of the 6th International Global WordNet, 2012, pp. 9-13.

22. B. Liu, X. Li, W. S. Lee and P. S. Yu. "Text classification by labeling words." in Proc. of the National Conference on Artificial Intelligence, Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press, 1999, pp. 425-430.

23. X. Ding and B. Liu."The utility of linguistic rules in opinion mining." in Proc. of the 30th annual international ACM SIGIR conference on Research and development in information retrieva, 2007, pp. 811-812.

24. T. Wilson, J. Wiebe and P. Hoffmann. "Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis." in Proc. of the conference on Human Language Technology and Empirical Methods in Natural Language Processing, Vancouver, British Columbia, Canada, Association for Computational Linguistics, 2005, pp. 347-354.

25. P. J. Stone, C. D. Dexter, S. S. Marshall and M. O. Daniel. "The General Inquirer: A Computer Approach to Content Analysis." Cambridge, MA: MIT Press, 1966.

26. J. W. Pennebaker,R. J. Booth and M. E. Francis. "Linguistic inquiry and word count: LIWC 2001." Mahway: Lawrence Erlbaum Associates, 2001.

27. R. Prabowo and M. Thelwall. "Sentiment analysis: A combined approach." J. Informetrics, vol. 3, 2009, pp. 143-157.

28. P. D. Turney. "Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews." in Proc. of the 40th Annual Meeting on Association for Computational Linguistics. Philadelphia, Pennsylvania, Association for Computational Linguistics, 2002, pp. 417-424.

29. M. Hu and B. Liu. "Mining and summarizing customer reviews." in Proc. of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining. Seattle, WA, USA, ACM, 2004, pp. 168-177.

30. J. Daudé, L. Daudé and G. Rigau. " Mapping WordNets Using Structural Information." in 38th Annual Meeting of the Association for Computational Linguistics (ACL'2000), Hong Kong, 2000.