Subjectivity and Sentiment Analysis of Arabic: A Survey

Mohammed Korayem[1,3], David Crandall[1], and Muhammad Abdul-Mageed[2]

[1] School of Informatics and Computing, Indiana University
Bloomington, Indiana USA
{mkorayem,djcran}@indiana.edu
[2] Department of Linguistics &
School of Library & Information Science,
Indiana University, Bloomington, USA
mabdulma@indiana.edu
[3] Department of Computer Science,
Fayoum University, Fayoum, Egypt

Abstract. Subjectivity and sentiment analysis (SSA) has recently gained considerable attention, but most of the resources and systems built so far are tailored to English and other Indo-European languages. The need for designing systems for other languages is increasing, especially as blogging and micro-blogging websites become popular throughout the world. This paper surveys different techniques for SSA for Arabic. After a brief synopsis about Arabic, we describe the main existing techniques and test corpora for Arabic SSA that have been introduced in the literature.

1 Introduction

The web has become a read-write platform where users are no longer strictly consumers of information but also producers. User-generated content, in the form of unstructured free text, is becoming an integral part of the web mainly because of the dramatic increase of social network sites, video sharing sites, online news, online reviews sites, online forums and blogs. Because of this proliferation of user-generated content, Web content mining is gaining considerable attention due to its importance for many businesses, governmental agencies and institutions. Subjectivity and sentiment analysis (SSA) is an important sub-area of Web content mining.

In natural language, *subjectivity* refers to aspects of language used to express opinions, feelings, evaluations, and speculations [14], including sentiment. The process of subjectivity classification refers to the task of classifying texts as either *objective* (e.g., *The new iPad was released*) or *subjective* (e.g., *The new iPad is cool*). Subjective text can further be classified by its *sentiment* and *polarity*. For sentiment classification, the task consists of identifying whether a subjective text is *positive* (e.g., *Egyptians inspired the world with their revolution!*), *negative* (e.g., *The bloodbaths in Syria are horrifying!*), *neutral* (e.g., *Obama may sign the bill*), or *mixed* (e.g., *The iPad is cool, but way too expensive*). These various types of SSA are gaining increasing attention because they provide an automated way to summarize vast quanitities of text (including reviewers, blogs, Twitter feeds, etc.) into the opinions that they express. This data is of

A. Ell Hassanien et al. (Eds.): AMLTA 2012, CCIS 322, pp. 128–139, 2012.

use to businesses and institutions who want to monitor feelings about their products and services. Private citizens can also benefit from this information to be able to compare sentiments about competing products, for example. Following a considerable body of related literature, we will henceforth use the terms *subjectivity and sentiment analysis* and *sentiment analysis* interchangeably.

Sentiment analysis can thus be viewed as a classification process that aims to determine whether a certain document or text was written to express a positive or a negative opinion about a certain object (e.g., a topic, product, or person). This process regards each document as a basic information unit. The process has been referred to as 'document level-sentiment classification' where the document is seen as an opinionated artifact. The more fine-grained problem of identifying the sentiment of every sentence has also been studied [30]. Sentiment analysis is typically performed using one of two basic approaches: *rule-based classifiers*, in which rules derived from linguistic study of a language are applied to sentiment analysis [15, 16, 18, 24], and *machine learning classifiers* in which statistical machine learning algorithms are used to learn signals of sentiment automatically [2, 26, 27, 30, 33, 34].

Currently, most of the systems built for sentiment analysis are tailored for the English language [26,30], but there has been some work on other languages. This paper reviews efforts to build SSA systems for Arabic, a Semitic language. After a brief discussion of the properties of Arabic in Section 2, we review resources and test corpora for Arabic in Section 3. Section 4 then reviews various approaches that have been employed for Arabic SSA.

2 Arabic

As the official language of 22 countries, Arabic is spoken by more than 300 million people, and is the fastest-growing language on the web (with an annual growth rate of 2,501.2% 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) [1]. There are about 65 million Arabic-speaking users online, or about 18.8% of the global Internet population [1].

Arabic is a Semitic language [35] and consists of many different regional dialects. While these dialects are true native language forms, they are typically used only in informal daily communication and are not standarized or taught in schools [22]. There is one formal written standard that is commonly used in written media and education throughout the Arab world called *Modern Standard Arabic* (MSA). There is a large degree of difference between MSA and most Arabic dialects, and, interestingly, MSA is not actually the native language of any Arabic country or group. MSA is syntactically, morphologically, and phonologically based on Classical Arabic (CA) [22], which is the language of the Qur'an (Islam's Holy Book).

Arabic has a very rich inflectional system and is considered one of the richest languages in terms of morphology [23]. Arabic sentential forms are divided into two types, nominal and verbal constructions [21]. In the verbal domain, Arabic has two word order patterns (i.e., Subject-Verb-Object and Verb-Subject-Object). In the nominal domain, a normal pattern would consist of two consecutive words, a noun (i.e., subject) then an adjective (subject descriptor).

3 Corpora and Lexicons

The availability of annotated corpora for training and testing is very important to enable progress on sentiment recognition systems. Collecting this data (and particularly the annotations) can be very labor-intensive. Fortunately, a number of research groups have developed and released Arabic sentiment analysis corpora, and we review this work here.

AWATIF is a multi-genre, multi-dialect corpus for Arabic SSA built by Abdul-Mageed and Diab [4, 7]. AWATIF is extracted from three different resources: the Penn Arabic Treebank (PATB), which is an existing collection of news wire stories in different domains (e.g. sports, politics, finance, etc.), Wikipedia user talk pages covering a variety of topics, and conversation threads from web forums on seven different sites. In annotating the corpus they used two different procedures, one that used untrained annotators via crowd sourcing technologies to give a coarse sentiment label (positive, negative, or neutral) to each sentence, and then one that used annotators trained with some linguistic background to label each sentence. The authors also manually created an adjective polarity lexicon, giving coarse labels (again positive, negative or neutral) to each of nearly 4,000 Arabic adjectives. In a related work, Abdul-Mageed and Diab [5] use a machine translation procedure to translate available English lexicons (e.g., from [20] and [9]) into Arabic. They retrieve 229,452 entries, including expressions commonly used in social media.

Opinion corpus for Arabic (OCA) is a corpus of text from movie review sites by Rushdi-Saleh *et al.* [34], and includes a parallel English version called EVOCA. The corpus consists of 500 reviews, half negative and half positive. The raw reviews contained a number of challenges which the authors attempted to fix manually, including filtering out spurious and unrelated comments, romanization of Arabic, multi-language reviews, differing spellings of proper names, and movie reviews that were more opinions of the cultural and political themes of a movie than the film itself. (As an example of the latter issue, the movie "Antichrist" has a rating of 6.7 in IMDB but a rating of 1 in the reviews on the Arabic blog.) OCA and EVOCA performed standard pre-processing on the corpus, including correcting spelling mistakes and deleting special characters, and also have made available unigram, bigram, and trigrams for the dataset.

MPQA subjective lexicon & Arabic opinion holder corpus: Another corpus for Arabic opinion holder and subjectivity lexicon is proposed by Elaranoty et al. [18], who created an Arabic news corpus. They crawled 150 MB of Arabic news and manually annotated 1 MB of the corpus for opinion holder. The opinion holder corpus was annotated by three different people. Any conflict emerging because of different annotations was solved using majority voting. For preprocessing the corpus, the Research and Development International (RDI) tool (http://www.rdi-eg.com) was used to handle the morphological analysis of Arabic sentences and assign parts of speech (POS) tags. Finally, semantic analysis of the words were done. Arabic Named Entity Recognition (ANER) [3] was used for extracting names from documents. The proposed Arabic subjectivity lexicon contains strong as well as weak subjective clues by manually translating the MPQA lexicon [37].

An **Arabic Lexicon for Business Reviews** was proposed by Elhawary and Elfeky [19]. Elhawary and Elfeky [19] used the similarity graph to build an Arabic lexicon.

The similarity graph is a type of graph where the two words or phrases have an edge if they are similar on polarity or meaning. The weight of the edge represents the degree of similarity between two nodes. Usually, this graph is built in an unsupervised manner based on lexical co-occurrences from large Web corpora. Here, the researchers initially used a small set of seeds and then performed label propagation on an Arabic similarity graph. For building the Arabic similarity graph, a list of seeds for positive, negative and neutral are used. The Arabic lexicon created from the similarity graph consists of two columns where the first column is the word or phrase and the second column represent the score of the word which is the sum of the scores of all edges connected to this node (word/phrase). They applied filtering rules to avoid both the sparseness of the data and garbage nodes. Garbage nodes caused the top 200 positive words to be non-positive. They removed nodes with a high number of weighted edges and kept the 25 top ranked synonyms of the word. The top 25 synonyms of positive words are 90% positive. This ratio became 50-60% when considering all synonyms. The sentiment of the review is computed based on the sentiment of the sentences. That is, sentence boundary detection is used, and negation is also used, to flip the sentiment score from positive to negative and vice versa. There are around 20 Arabic words for negation. Sentences greater than 120 character are ignored. The results show that the created Arabic lexicon has high precision but has low recall.

Another subjectivity lexicon is proposed by El-Halees [17]. This lexicon is built manually based on two resources, the SentiStrength project and an online dictionary. They translated the English list from SentiStrength project and then manually filtered it. Common Arabic words were added to the lexicon.

4 Subjectivity and Sentiment Analysis Systems and Methods for Arabic

We now turn to reviewing the different methods for Arabic subjectivity and sentiment analysis that have been proposed in the literature. First, we describe language-independent feature selection/extraction methods applied to Arabic. Then, we discuss the systems employing standard IR methods (e.g., TF*IDF) and hybrid classifiers. Finally, we discuss systems employing Arabic-specific features and those tailored for extracting opinion holders.

4.1 Language-Independent Feature Selection and Extraction

One way of building sentiment analyses for languages other than English or building systems that work for multiple languages is to extract and select features that do not depend on the language itself. Different approaches have been followed to select and extract these features, including: (1) Entropy Weighted Genetic algorithms [2], (2) Feature Subsumption [38], (3) Local grammar-based methods [11, 12], (4) Positional features [32] and (5) Common seeds word methods [25]. Among these methods, we are aware of only two that have been applied to Arabic: entropy weighted genetic algorithms and local grammar methods.

Abbasi *et al.* [2] used genetic algorithms to select language features for both Arabic and English. Genetic Algorithms (GA) are a general class of techniques that apply the concept of evolution to general optimization problems. Entropy Weighted Genetic Algorithms (EWGA) combine genetic algorithms with information gain (IG) to perform the feature selection. In particular, IG is used to select the initial set of features for the initial stage of the GA, and is also applied during the cross-over and mutation stages. EWGA is applied to select features for sentiment analysis in a corpus of Web forum data containing multiple languages [2]. They used two types of features, stylistic features and lexical features. They avoided semantic features because they are language dependent and need lexicon resources, while the limitation of their data prevent using of linking features.

The paper evaluates the proposed system on a benchmark testbed consisting of 1000 positive and 1000 negative movie reviews. Their system that uses feature selection outperforms several previous systems [28, 29, 31, 36]. Using this system, they achieved an accuracy rate of 91% while other systems achieved accuracy rates between 87-90% on the movie review data set. They were also able to achieve 92% accuracy on Middle Eastern forums and 90% on US forums using the EWGA feature selection method.

Use of Local Grammar is another method that can be used to extract sentiment features [11]. Ahmed *et al.* [12, 13] applied this approach to documents from the financial news domain. They identified domain-specific key words by looking for words that occurred often in a corpus of financial news but relatively infrequently in a general corpus. Using the context around these words they built a local grammar to extract sentiment-bearing phrases. They applied their approach to Arabic, English and Chinese. They evaluated the system manually and achieved accuracy rates between 60-75% for extracting the sentiment bearing phrases. Importantly, the proposed system could be used to extract the sentiment phrases in financial domain for any language.

4.2 Standard IR and Hybrid Classifiers

Here we describe systems employing standard IR methods (e.g., TF*IDF) and hybrid classifiers.

The work of Rushdi *et al.* [33, 34] builds machine learning classifiers using both the OCA and EVOCA corpora. They use both Support Vector Machines (SVMs) and Naive Bayes (NB) classifiers, reporting 90% F-measure on OCA and 86.9% on EVOCA using SVMs. They show that SVMs outperform the NB classifier, which is common in text classification tasks. Results also show that there is no difference between using term frequency (tf) and the slightly more complicated term frequency-inverse document frequency (tf-idf) as weighting schemes. Experiments also show that stemming words before feature extraction and classification nearly always degrades the results.

Elhawary and Elfeky [19] present a system for sentiment analysis on Arabic business reviews, with the specific goal of building a web search engine that would automatically annotate returned pages with sentiment scores. The system has several components. The first component classifies whether an Internet page is a review or not. For this component, they extend an in-house multi-label classifier to work for Arabic such that its task is to assign a tag from the set {*review, forum, blog, news, shopping store*} to a document. To build an Arabic review classifier data set, 2000 URLs were collected and

more than 40% of them were found to be reviews through manual labeling. This data set was collected by searching the web using keywords that usually exist in reviews (such as "the camera is very bad"). The authors translated the lists of keywords collected and add to them a list of Arabic keywords that usually appear in opinionated Arabic text. The final list contained 1500 features and was used to build an AdaBoost classifier, using 80% of the data for training and the rest for testing. After a document is classified as belonging to the Arabic review class or not, a second component of the system analyzes the document for its sentiment. They build an Arabic lexicon based on a similarity graph for use with the sentiment component. The final component of the system is designed to provide the search engine with an estimate of the sentiment score assigned to a document during the search.

A combined classification approach is proposed by El-Halees [17] for document level sentiment classification. He applied different classifiers in a sequential manner. A lexicon-based classifier is first used to estimate the sentiment of a document based on an aggregation of all the opinion words and phrases in the document. However, because some documents lack enough opinion words to use this lexicon-based classifier, a second phase uses a maximum entropy classifier. All classified documents from first classifier are used as the training set for this maximum entropy classifier, which is then used to compute the probability that a given document belongs to a certain sentiment class. In particular, if the probability is greater than a threshold of 0.75, then the document is assigned a class, and otherwise the document is passed to the next stage. The final stage is a *k*-nearest neighbors (KNN) classifier that finds the nearest neighbors for the unannotated document using the training set coming from the previous two classifiers.

The corpus used for evaluation consisted of 1134 documents collected from different domains (e.g., education, politics, and sports), with 635 positive documents (with 4375 positive sentences) and 508 negative documents (with 4118 negative sentences). Preprocessing is applied to filter HTML tags and non-textual contents are removed. Alphabets are normalized and some misspelled words are corrected. Sentences are tokenized, stop words are removed, and an Arabic light stemmer is used for stemming the words, and *TF-IDF* is used for term weighting. The paper reports an f-measure of 81.70% averaged over all domains for positive documents and 78.09% *F*-measure for negative documents. The best *F*-measure is obtained in the education domain (85.57% for the positive class and 82.86% for the negative class).

4.3 Arabic-Specific, Social Media, and Genre-Specific Features

Other techniques use the linguistic features of Arabic in order to perform sentiment analysis, by analyzing the grammatical structure of Arabic [21] and Arabic-specific morphological features [4, 6–8].

Farra *et al.* [21] proposed Arabic sentence-level sentiment classification, considering two different approaches: a grammatical approach and a semantic approach. The grammatical approach is based on Arabic grammatical structure and combines the verbal and nominal sentence structures in one general form based on the idea of actor/action. In this approach, the subjects in verbal and nominal sentences are actors and verbs are actions. They manually label action and actor tags in sentences and used these tags

as features. Their feature vector constitutes the following dimensions: sentence type (verbal or nominal), actor, action, object, adjective, type of pronoun and noun, transition (the type of word linking the current sentence with the previous sentence), word polarity (positive, negative, neutral) and sentence class.

The second approach proposed by Farra *et al.* [21] combined syntactic and semantic features by extracting some features like the frequency of positive, negative, and neutral words, the frequency of special characters (e.g., "!"), the frequency of emphasis words (e.g., "really" and "especially"), the frequency of conclusive and contradiction words, etc. For extracting the semantics of the words, the paper builds a semantic interactive learning dictionary which stores the semantic polarity of word roots extracted by stemmer. The system asks the user for the polarity of a word if it has not yet been learned.

For evaluation of the grammatical approach, only 29 sentences are annotated manually with part-of-speech tags. They report 89.3% accuracy using an SVM classifier with 10-fold cross validation. Sentences from 44 random documents are used for evaluating the semantic and syntactic approach using a J48 decision tree classifier. They report 80% accuracy when the semantic orientation of the words extracted and assigned manually is used, and 62% when the dictionary is used. They also classified the documents by using all sentence features and chunking the document into different parts, reporting 87% accuracy with an SVM classifier when documents divided into 4 chunks and neutral class excluded.

Abdul-Mageed et al. [4, 6–8] created sentence-level annotated Arabic corpora and built subjectivity and sentiment analysis systems exploiting them. In their systems these authors use various types of features, including language independent features, Arabic-specific morphological features, and genre-specific features. In [6,8], they classify MSA news data at the sentence level for both subjectivity and sentiment. They use a two-stage SVM classifier, where a subjectivity classifier is first used to separate subjective from objective sentence. In a second stage, subjective sentences are classified into positive and negative cases, with an assumption that neutral sentences will be treated in a future system. These authors make use of both language-independent and Arabic-specific features. The language independent features include a feature indicating the domain of the document (e.g., politics, sports) from which a sentence is derived, a *unique* feature where all words occuring less than four times are replaced by the token "UNIQUE", *N-gram* features including all *N*-grams of frequency less than 4, and an *adjective* feature for adjectives indicating the occurrence of a polarized adjective based on a pre-developed polarity lexicon of 3982 entries. Classification accuracy of 95.52% are reported. The results showed that the adjective feature is very important, as it improved the accuracy by more than 20% and the unique and domain features are also helpful.

Other work by Abdul-Mageed *et al.* [10] presents SAMAR, an SVM-based system for Subjectivity and Sentiment Analysis (SSA) for Arabic social media genres. They tackle a number of research questions, including how to best represent lexical information, whether standard features are useful, how to treat Arabic dialects, and whether genre specific features have a measurable impact on performance. The authors exploit data from four social media genres: Wikipedia Talk Pages, Web forums, chat, and Twitter tweets. The corpus includes data in both Modern Standard Arabic and dialectal

Arabic. These authors break down their data into 80% training, 10% development, and 10% testing and exploit standard SSA features (e.g., the "unique" feature, a wide coverage polarity lexicon), social and genre features (e.g., the gender of a user), and a binary feature indicating whether a sentence is in MSA or dialectal Arabic. They are able to significantly beat their majority class baselines with most data sets. The results suggest that they need individualized solutions for each domain and task, but that lemmatization is a feature in all the best approaches.

Table 1 summarizes the SSA systems which are described above.

Table 1. Summary of different Arabic SSA systems

System	Type	Features	SSA level	Corpus	Advantages	Disadvantages
[2]	ML	Stylistic + LF	Doc	Movie reviews, web forums	– Language independence – Effective feature selection	– High computational cost
[12, 13]	NC	domain-specific lexical features	Phr	Financial news	– Simple method – Language independence	– No sentiment classification (only phrase extraction)
[33, 34]	ML	LF	Doc	Web reviews	– Simple features – Introduces OCA corpus	– No Arabic-specific features
[19]	ML	LF	Doc	Business reviews	– Builds large-scale lexicon – Computes soft sentiment score (in addition to hard classification)	– No Arabic-specific features
[17]	LC+ ML	LF	Doc	Multi-domain	– Combines lexical and ML – Multi-domain	– No Arabic-specific features
[21]	ML	Syntactic & LF	Sen+ Doc	News	– Combines LF & syntactic	– Evaluated on small dataset
[4, 6–8]	ML	LF, syntactic & genre-specific, social media features	Sen	News, social media	– Combines language-independent and Arabic-specific features – Incorporates dialectal Arabic – Employs a wide-coverage polarity lexicon	–Some genre and social media features are costly to acquire

Legend

Classification types: ML=Machine Learning, CL=Rule or lexicon-based classifiers, NC=No classification.

Features: LF=Lexical features.

SSA level: Doc=Document-level, Phr=Phrase-level, Sen=Sentence-level classification.

4.4 Opinion Holder Extraction

Different approaches for extracting the opinion holder in Arabic are proposed in [18]. Their approach is based on both pattern matching and machine learning. They extract three different types of opinion holders. The first type of opinion holder is opinion holder for speech events, which is defined as a subjective statement said directly by someone or claimed to be said by someone. In this way, they combine the direct speech

event and indirect speech event in this type. The second type of opinion holder is defined as related to an opinion holder that expresses sentiment towards certain opinion subject. The third type is defined as related to expressive subjective elements (e.g., emotions, sarcasm) expressed implicitly. Definitely the third type is the hardest type to extract because it depends on the meaning of the words rather than the structures. The first approach used in [18] to extract opinion holders is based on pattern matching. They manually extract 43 patterns where the morphological inflections of the words are neglected. Examples of these patterns are "And <holder> expressed his objection about ....", or "And adds <holder>...." A pattern-based opinion holder classifier is built using the extracted patterns. The following rule to extracting an opinion holder is followed: the opinion holder is retrieved if it contains a subjective statement or a named entity and it contains a statement that is classified as objective or subjective using a high-precision classifier.

While the first approach is based on pattern matching, the second and third approaches are based on machine learning. The authors formulate the opinion holder problem as a classification problem where each word in the corpus is classified as "Begining of a holder (B-holder)", "Inside a holder (I-holder)" or "Non holder". A conditional random field (CRF) probabilistic discriminative model is used for classification. The authors build the CRF classifier based on a set of lexical, morphological, and semantic features. Pattern matching is used as a source for additional features for training the classifier in the third approach. Syntactic features are not used because of a lack of a robust general Arabic parser. The lexical features used are the focus word itself and window of size 3 around it (i.e., previous and next three words). The second type of features, i.e., semantic field features, are generated by grouping the semantically related words and giving them the same probability. In that way the handling of a missing word of the group in training data will not affect the performance if any word of the group appeared in the test data. The third feature type used is POS Tags generated by the RDI morphological analyzer. The set of tags generated by the RDI analyzer is reduced to a small set of tags and this small set are used as features. In addition, base phrase chunk and named entity recognition features are used. Finally, a feature based on pattern matching is used such that it is detected whether any word is part of the patterns extracted manually in the first approach or not.

Experimental results on the Arabic Opining Holder corpus show that machine learning approaches based on CRF achieve better results than the pattern matching approach. The authors report 85.52% precision, 39.49% recall, and 54.03% F-measure. Authors justify the performance degradation of the system by stating that it is due to the lower performance of Arabic NLP tools compared to those of English as well as the absence of a lexical parser.

5 Conclusion

To conclude, we surveyed the different methods for building subjectivity and sentiment analysis systems for Arabic. The available resources for Arabic sentiment analysis are introduced. Here, it is suggested that a nuanced method to be followed in building a sentiment analysis system for Arabic employs not only language-independent, but

also Arabic-specific features; exploits wide-scale, domain-specific polarity lexicons; and leverages genre- and social media-specific features. While it is costly to build resources tailored to Arabic and acquire certain types of features (e.g., genre- and social media-specific features), this route yields high performance and brings interesting insights to the classification task. Alternatives to this method would be transferring the sentiment knowledge from English into the Arabic language or to use language independent methods.

References

1. Internet world stats, http://www.internetworldstats.com/stats7.htm
2. Abbasi, A., Chen, H., Salem, A.: Sentiment analysis in multiple languages: Feature selection for opinion classification in Web forums. ACM Transactions on Information Systems (TOIS) 26(3), 12 (2008)
3. AbdelRahman, S., Elarnaoty, M., Magdy, M., Fahmy, A.: Integrated machine learning techniques for arabic named entity recognition. International Journal of Computer Science Issues, IJCSI 7(4), 27–36 (2010)
4. Abdul-Mageed, M., Diab, M.: AWATIF: A Multi-Genre Corpus for Modern Standard Arabic Subjectivity and Sentiment Analysis. In: Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC 2012). European Language Resources Association, ELRA (2012)
5. Abdul-Mageed, M., Diab, M.: Toward building a large-scale Arabic sentiment lexicon. In: Proceedings of the 6th International Global Word-Net Conference, Matsue, Japan (2012)
6. Abdul-Mageed, M., Diab, M., Korayem, M.: Subjectivity and sentiment analysis of modern standard Arabic. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, vol. 2, pp. 587–591. Association for Computational Linguistics (2011) (short papers)
7. Abdul-Mageed, M., Diab, M.T.: Subjectivity and sentiment annotation of modern standard arabic newswire. In: Proceedings of the 5th Linguistic Annotation Workshop, LAW V 2011, pp. 110–118 (2011)
8. Abdul-Mageed, M., Korayem, M.: Automatic Identification of Subjectivity in Morphologically Rich Languages: The Case of Arabic. In: Proceedings of the 1st Workshop on Computational Approaches to Subjectivity and Sentiment Analysis (WASSA), pp. 2–6 (2010)
9. Abdul-Mageed, M., Korayem, M., YoussefAgha, A.: Yes we can?: Subjectivity annotation and tagging for the health domain. In: Proceedings of the International Conference Recent Advances in Natural Language Processing, RANLP, Hissar, Bulgaria (2011)
10. Abdul-Mageed, M., Kuebler, S., Diab, M.: 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, pp. 19–28. Association for Computational Linguistics (2012)
11. Agić, Ž., Ljubešić, N., Tadić, M.: Towards sentiment analysis of financial texts in croatian. In: Proceedings of the Seventh International Conference on Language Resources and Evaluation, LREC 2010 (2010)
12. Ahmad, K., Cheng, D., Almas, Y.: Multi-lingual sentiment analysis of financial news streams. In: Proceedings of the 1st International Conference on Grid in Finance (2006)
13. Almas, Y., Ahmad, K.: A note on extracting sentiments in financial news in English, Arabic & Urdu. In: Proceedings of Workshop on Computational Approaches to Arabic Script-based Languages (2007)

14. Banfield, A.: Unspeakable sentences: Narration and representation in the language of fiction. Routledge & Kegan Paul Boston (1982)
15. Brooke, J., Tofiloski, M., Taboada, M.: Cross-linguistic sentiment analysis: From English to Spanish. In: Proceedings of the 7th International Conference on Recent Advances in Natural Language Processing, Borovets, Bulgaria, pp. 50–54 (2009)
16. Denecke, K.: Using SentiWordNet for multilingual sentiment analysis. In: IEEE 24th International Conference on Data Engineering Workshop, ICDEW, pp. 507–512. IEEE (2008)
17. El-Halees, A.: Arabic Opinion Mining Using Combined Classification Approach. In: Proceedings of the International Arab Conference on Information Technology, ACIT (2011)
18. Elarnaoty, M., AbdelRahman, S., Fahmy, A.: A Machine Learning Approach For Opinion Holder Extraction Arabic Language. CoRR, abs/1206.1011 (2012)
19. Elhawary, M., Elfeky, M.: Mining Arabic Business Reviews. In: Proceedings of International Conference on Data Mining Workshops (ICDMW), pp. 1108–1113. IEEE (2010)
20. Esuli, A., Sebastiani, F.: Sentiwordnet: A publicly available lexical resource for opinion mining. In: Proceedings of the 5th Conference on Language Resources and Evaluation (LREC 2006), vol. 6, pp. 417–422. Citeseer (2006)
21. Farra, N., Challita, E., Assi, R., Hajj, H.: Sentence-Level and Document-Level Sentiment Mining for Arabic Texts. In: Proceedings of International Conference on Data Mining Workshops (ICDMW), pp. 1114–1119. IEEE (2010)
22. Habash, N.: Introduction to Arabic natural language processing. Synthesis Lectures on Human Language Technologies 3(1), 1–187 (2010)
23. Habash, N., Rambow, O., Roth, R.: Mada+ tokan: A toolkit for arabic tokenization, diacritization, morphological disambiguation, POS tagging, stemming and lemmatization. In: Proceedings of the 2nd International Conference on Arabic Language Resources and Tools (MEDAR), pp. 242–245 (2009)
24. Kim, J., Jung, H., Lee, Y., Lee, J.: Conveying Subjectivity of a Lexicon of One Language into Another Using a Bilingual Dictionary and a Link Analysis Algorithm. International Journal of Computer Processing of Oriental Languages 22(2), 205–218 (2009)
25. Lin, Z., Tan, S., Cheng, X.: Language-independent sentiment classification using three common words. In: Proceedings of the 20th ACM International Conference on Information and Knowledge Management, pp. 1041–1046. ACM (2011)
26. Liu, B.: Sentiment analysis and subjectivity. In: Handbook of Natural Language Processing, pp. 627–666 (2010)
27. Mihalcea, R., Banea, C., Wiebe, J.: Learning multilingual subjective language via cross-lingual projections. In: Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics, vol. 45, pp. 976–983. Association for Computational Linguistics (2007)
28. Mullen, T., Collier, N.: Sentiment analysis using support vector machines with diverse information sources. In: Proceedings of Conference on Empirical Methods in Natural Language Processing, EMNLP, vol. 4, pp. 412–418 (2004)
29. Pang, B., Lee, L.: A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In: Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics, p. 271. Association for Computational Linguistics (2004)
30. Pang, B., Lee, L.: Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval 2(1-2), 1–135 (2008)
31. Pang, B., Lee, L., Vaithyanathan, S.: Thumbs up?: Sentiment classification using machine learning techniques. In: Proceedings of Conference on Empirical Methods in Natural Language Processing, pp. 79–86. Association for Computational Linguistics (2002)
32. Raychev, V., Nakov, P.: Language-independent sentiment analysis using subjectivity and positional information. In: Proceedings of the International Conference Recent Advances in Natural Language Processing, RANLP, pp. 360–364 (2009)

33. Rushdi-Saleh, M., Martín-Valdivia, M., Ureña-López, L., Perea-Ortega, J.: Bilingual Experiments with an Arabic-English Corpus for Opinion Mining, pp. 740–745 (2011)
34. Rushdi-Saleh, M., Martín-Valdivia, M., Ureña-López, L., Perea-Ortega, J.: Oca: Opinion corpus for Arabic. Journal of the American Society for Information Science and Technology 62(10), 2045–2054 (2011)
35. Versteegh, K., Versteegh, C.: The Arabic Language. Columbia University Press (1997)
36. Whitelaw, C., Garg, N., Argamon, S.: Using appraisal groups for sentiment analysis. In: Proceedings of the 14th ACM International Conference on Information and Knowledge Management, pp. 625–631. ACM (2005)
37. Wilson, T., Wiebe, J., Hoffmann, P.: Recognizing contextual polarity in phrase-level sentiment analysis. In: Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing, pp. 347–354. Association for Computational Linguistics (2005)
38. Zhai, Z., Xu, H., Li, J., Jia, P.: Feature Subsumption for Sentiment Classification in Multiple Languages. In: Zaki, M.J., Yu, J.X., Ravindran, B., Pudi, V. (eds.) PAKDD 2010, Part II. LNCS, vol. 6119, pp. 261–271. Springer, Heidelberg (2010)