Lexicon-based Sentiment Analysis for Persian Text

Fatemeh Amiri[1], **Simon Scerri**[2] **and Mohammad H. Khodashahi**[3]
[1]Media Informatics, RWTH Aachen University, Germany
[2]Fraunhofer IAIS, Sankt Augustin, NRW, Germany
[3]Institute of Computer Science, University of Bonn, Germany
amirii.fatemeh@gmail.com
simon.scerri@iais.fraunhofer.de
mh.khodashahi@gmail.com

Abstract

The vast information related to products and services available online, of both objective and subjective nature, can be used to provide contextualized suggestions and guidance to possible new customers. User feedback and comments left on different shopping websites, portals and social media have become a valuable resource, and text analysis methods have become an invaluable tool to process this kind of data. A lot of business use-cases have applied sentiment analysis in order to gauge people's response to a service or product, or to support customers with reaching a decision when choosing such a product. Although methods and techniques in this area abound, the majority only address a handful of natural languages at best. In this paper, we describe a lexicon-based sentiment analysis method designed around the Persian language. An evaluation of the developed GATE pipeline shows an encouraging overall accuracy of up to 69%.

1 Introduction

In comparison to other more popular and widespread language, few research efforts have sought to provide text analytics services targeting Persian text documents on the Web. As the official language of Iran, Afghanistan, and Tajikistan and an estimated 110 million people, we feel that the Persian language has not been given the attention it deserves. Besides attaining merit from a purely linguistic point of view, providing technologies for Persian text analysis has also business implications in the regions where the language remains a preferred working language. In particular, sentiment analysis has a high potential in providing insights for several Persian online communities and social media. Most of the limited available techniques have employed Machine Learning (ML) algorithms, such as Support Vector Machine-based (SVM) methods. In contrast, our approach is based on a manually-created lexicon enriched with sentiment scores; coupled with hand-coded grammar rules. In tackling our objective, we are faced with language-specific challenges and constraints. In the Persian language there is typically a large difference between formal and informal writing styles. There is also a high level of complexity due to the frequent morphological operations. Besides a complex morphology, Persian has some other distinctive features, such as lexicon intricacy, a high context sensitivity of the script, and a free words order due to independent case-marking (Hajmohammadi and Ibrahim, 2013). Therefore models used in approaches behind other languages, or even aspects of which, can hardly be used in Persian text analytics methods.

In this paper, we describe how we approached the language-specific challenges when designing and implementing a lexicon-based sentiment analysis method for Persian text. An evaluation of this method is also presented. But before we provide an overview of related work in this area.

2 Related Work

As a technique, sentiment analysis has improved significantly in recent years, especially for mainstream languages such as English. The technique has an especially important role in business and financial circles. Efforts such as (Feldman et al., 2011) have specifically focused on stock markets and market predictions, whereas others focused on deriving changing opinions and perceptions from subjective information shared on social networks (Pak and Paroubek, 2010). Many studies have been performed to try and identify a

Proceedings of Recent Advances in Natural Language Processing, pages 9–16,
Hissar, Bulgaria, Sep 7–9 2015.

superior approach in the many techniques available (Feldman, 2013), in order to attain better results and higher accuracies. Different surveys have been carried out, with different viewpoints and results (Liu, 2012) (Liu, 2010). A large share of sentiment analysis techniques employ learning-based approaches (Pang et al., 2002) (Jo and Oh, 2011). Of these the most promising are SVM- and Nave Bayes-based methods. Using a supervised classification task, these methods attain up to 82.9% accuracy (Hajmohammadi and Ibrahim, 2013). However, various drawbacks have been noted, such as their strict reliance on a corpus of human-coded texts for training, and their domain dependency (Basiri et al., 2014) (Taboada et al., 2011).

A contrasting approach is the use of lexicon-based methods (Ding et al., 2008) (Thelwall et al., 2010), which calculate a documents orientation from the semantic orientation of words or phrases within that document (Turney, 2002). Sentiment-bearing words and phrases forming a sentiment lexicon (Liu, 2012) can be derived from different resources. Some have employed seed words to expand the final list of words (Hatzivassiloglou and McKeown, 1997), or use existing linguistic resources like the ANEW words (Bradley and Lang, 1999), SentiWordNet (Baccianella et al., 2010) and WordNet Affect (Strapparava et al., 2004).

Some research efforts have satisfactorily mixed the two above approaches to gain a better response (Mudinas et al., 2012). Although future work will consider extending our method with aspects from the first of the two approaches, for the moment we have opted to investigate a technique based solely on the second approach. Other surveyed research efforts, including the ones cited above, have already provided similar techniques that identify the orientation of a document based on the polarity of adjectives in a dictionary. However, they addressed either English (Hatzivassiloglou and McKeown, 1997) or other languages such as Urdu (Syed et al., 2010), Chinese (Zagibalov and Carroll, 2008), French (Ghorbel and Jacot, 2011) or Arabic (Abdul-Mageed et al., 2011).

Of the surveyed efforts which tackle the Persian language, a majority also utilized machine learning approaches. Bagheri and Saraee (Saraee and Bagheri, 2013) devised a learning-based approach that employs Nave-Bayes text classification. They proposed a new feature selected method (MMI) and reported a performance of 70%. Hajmohammadi and Ibrahim (Hajmohammadi and Ibrahim, 2013) used standard machine learning techniques incorporated into the domain of online Persian-written movie reviews to automatically classify reviews as either positive or negative. They also combined Nave-Bayes and SVM, in conjunction with six feature presentations concerning n-gram presence/frequency in order to examine the effects of the classifiers and the feature options on Persian sentiment classification.

More recently, a lexicon-based unsupervised approach (Basiri et al., 2014) addressed specific Persian text processing difficulties, such as different forms of writing styles and ignoring short spaces between words in texts. The approach utilises the SentiStrength library, which applies a combined method to detect the polarity and strength of short informal social texts. However, as this library was designed around the English language, the authors rely on the translation of the core resulting list to Persian. The reported results indicate an F-measure of around 90%.

The major difference between our approach and the above-mentioned effort is that we use an own-constructed lexicon and involve a number of human annotators to provide multiple sentiment scores. In resolving any resulting conflicts, we also address the issue of subjectivity. Therefore, our approach is in theory more appropriate as the generated lexicon and polarity pairs are Persian language-specific, whereas language translations such as the method used in the above-mentioned approach are problematic since languages are intrinsically different.

Our final aim is to outperform existing ML-based methods and achieve an acceptable F-measure. The evaluation results of this approach will then indicate whether our approach has any value, so that a more comprehensive effort at collecting key-word/phrase and polarity pairs will result in an improved approach that has the potential to rival the results reported by Basiri et. al.

3 Approach

3.1 Data Collection

For our lexicon-based sentiment analysis technique we needed a wide range of Persian vocabulary entries, and their sentiment. As no Persian API was available for achieving this requirement, we opted to manually gather a number of Persian adjectives, words and expressions (7179) from two online Persian language resources[1] . The criteria for selecting these gazetteer entries, as followed by the two native speakers authoring this paper, were the following:

- Terms (words or multi-word expressions) that can alter or influence the sentiment of a given statement in any conceivable context.
- Gathered lexicons are used in either formal or informal communication between Persian people.
- Gathered lexicons correspond to either standard Persian or obsolete Persian as used by certain sections of native speakers.

As already mentioned the formal and informal styles of Persian writing has a huge impact on the semantics. In many cases one cannot understand the meaning of an informal textual comment unless they are a native speaker. So the need to enrich the lexicon with as many informal expressions and comments was as necessary, if not more pressing than, gathering all the formal forms. In addition, some of the collected words and adjectives correspond to the old usage of the language among older native speakers. Although these are not used regularly in daily speech or text, they are still important to make our gazetteer as varied and as broad as possible. The resulting terms have been saved in a personal database in preparation for the sentiment annotation phase described below.

3.2 Sentiment Annotation

The results of the collection process were stored in a database, and in order to achieve the required lexicon we then required to annotate each entry with a sentiment score. To support with this task, we set up a Web interface[2] that enables native Speakers to manually assign a score to random entries. At each click, the interface presented a new adjective which could then be voted either as having either a positive, negative or neutral sentiment expression. A five-tier scoring spectrum was considered but eventually discarded in favour of the three-tier option above, for the sole reason that it was cognitively easier for the volunteers to decide on an outcome, and as a result, more votes were expected.

The exercise was shared between a number of volunteers , following requests via own and extended social networks of a personal and academic nature. Half of the targeted volunteers were Persian students. As a result, the annotation was performed by people having different levels of education, age groups and sectors corresponding to the Persian society. For the 7179 adjectives in the database, we received a total of 8278 votes. This discrepancy is intended and is due to the decision to allow multiple voting by different volunteers. In cases where the opinion expressed contrasted, manual conflict resolution was performed following a discussion, or the inconclusive entry was marked as neutral. Future work can focus on these entries and flag their polarity as highly contextual.

3.3 A lexicon-based Sentiment Analysis Pipeline

Following the establishment of an annotated Persian sentiment lexicon, we designed and developed a linguistic pipeline based on the GATE framework (Cunningham et al., 2002). The pipeline utilizes existing components that were already available[3], namely a Persian tokenizer, sentence splitter and POS tagger. In addition, our lexicon was provided as the basis for the gazetteer, and JAPE (Cunningham et al., 1999) grammar rules were then manually coded to address the most general features of the Persian language in its written form. The pipeline and its components is depicted in Fig. 1. A breakdown of all these components is provided below.

3.3.1 Tokenizer

The imported tokenizer splits the text into very simple tokens like words, numbers, spaces and punctuation. As the Persian script is not case-sensitive like most Latin scripts, the employed tokenizer excludes similar checks.

3.3.2 Sentence Splitter

The imported sentence splitter fragments the text into sentences. It uses a list of abbreviations to

[1] We collected Persian adjectives from the Wiktionary open source dictionary: http://goo.gl/o0J8K0 and from a reference database for the Persian lanaguage at: http://dadegan.ir/

[2] http://www.computerssl.com/sentiment/

[3] Although the library and components imported in our pipeline have not been made available online, they were kindly supplied by the author: http://sazvar.student.um.ac.ir/

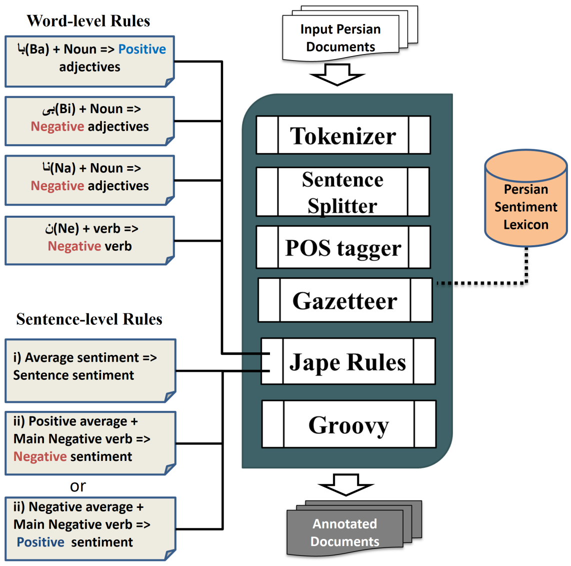

Figure 1: The Sentiment Analysis Pipeline

help distinguish sentence-marking full stops from other kind of splits.

3.3.3 POS Tagger

The imported tagger produces a part-of-speech tag as an annotation for each word or symbol. It uses a default lexicon and rule set which can be manually modied.

3.3.4 Gazetteer

The gazetteer includes all information resulting from the data collection and sentiment analysis exercises. In short, the employed gazetter is the basis for our lexicon-based approach. Whenever a gazetteer entry appears in the text, it is marked and assigned a sentiment score accordingly.

3.3.5 Hand-coded Persian grammar patterns

JAPE provides finite state transduction over annotations based on regular expressions. In our pipeline, we utilize JAPE rules to identify regular expressions we have formulated as a grammar base for Persian. Therefore, together with the gazetteer, this is one of the main contributions presented in this paper. We designed rules in two phases:

1. Phase I: patterns are focussed on and around each individual text-based token (i.e. words) in an input text segment.

2. Phase II: we address the sentiment of the entire text segment, based on the computed sentiment of each individual word.

Both phases are also depicted in Fig. 1. To identify the sentiment at the word-level, we created rules to consider an alternate sentiment to that otherwise identified by the gazetteer due to a special prefix and postfix. For example, in Persian, in a majority of cases a "Ba" prefix before a noun alters the polarity to positive, whereas a "Bi" or "Na" prefix alters it to negative. Some examples of the above alterations are shown in the table below. Similarly, we have catered for the linguistic alternative of verbs. Most notably, in Persian the verbs can be given a negative connotation by using "n" as a prefix (equivalent to the effect of having a do not before a verb in English). Examples are also shown in the below table.

Persian (before)	Persian (after)	English (before)	English (after)
اخلاق	بی اخلاق	Moral	Immoral
ادب	بی ادب	Politeness	Impolite
معرفت	بامعرفت	Wisdom	Wise
درست	نادرست	Correct	Incorrect
ندارد		Don't understand	
نمیفهمد		Don't have	

In many cases, in order to calculate the sentiment of an entire sentence or text segment it is not simply a case of averaging or combining the sentiment of each word as identified in Phase I. Some adjectives or phrases have a direct effect on the entire sentence, e.g., the presence of just one special negative verb in a sentence that otherwise consists of mostly positive words, alters the entire polarity of the sentence to negative (irony). Therefore, in this second phase the JAPE rules follow this sequence:

1. Step 1: the number of positive and negative words in a sentence are counted and the average is used to identify the polarity of the sentence

2. Step 2: the main verb of the sentence is identified, and if it matches one of the known exceptional negative verbs, the polarity of the pre-computed sentence is reversed

Examples of cases which are addressed by step 2 above are in the table below, with their English language equivalent.

Persian	English Equivalent
"این فیلم بازیگران معروف زیادی داشت ولی نتوانست نظر مردم را جلب کند"	"That film had a lot of famous actors but it couldn't attract people's attention."
"فیلم خوبی نبود"	"It wasn't a good film!"
"آدم دروغگویی نیست"	"He is not a liar"

3.3.6 Groovy scripting processing resource

The result of the two JAPE phases are then forwarded to the Groovy scripting processing resource, for which GATE also provides support. The Groovy plugin is used to count the number of positive and negative annotations in a given piece of text and determine an overall polarity score. Therefore, this can also be considered a third phase in the sentiment analysis, which takes place at the paragraph or entire document level. It must be noted that at the moment, the final sentiment score determined is either positive, negative or neutral.

4 Evaluation

In order to evaluate the performance of our approach, we performed two experiments. In the initial one, we relied on a pre-existing corpus of annotated text, based on the availability of reviews related to accommodation online. However, the information available here was not in a form to enable us to confidently reach conclusive results. Therefore, in a second experiment, we again instructed native speakers to rate a large amount of Persian news items and compared their judgment against the ones determined by our pipeline. Details and results are presented below.

4.1 Corpus-based Evaluation

In this experiment we choose customer reviews that are available online for a website[4] specializing in hotel reservation and accommodation in different cities of Iran. Although its popularity has recently seen a downturn[5], the site has been used for 15 years and therefore there are a lot of valuable reviews that can be used for this kind of experiment. Website visitors are able to leave their opinions about their previous experience in a hotel (including references to price, quality and local sightseeing) by filling verifiable identification fields, thus meaning that the expressed opinions are probably genuine and reliable. The main problem with this corpus is that the reviews are star base, on a scale of 1 (poor) to 5 (excellent) stars. Therefore, in order to be able to compare to the results generated by the developed pipeline we were required to map this expression of sentiment as follows:

- 1 and 2 stars: Negative
- 3 stars: Neutral
- 4 and 5 stars: Positive

From the above, we generated a corpus of test and evaluation data. The reviews were each passed on to the pipeline, and the calculated sentiment score was directly compared to the ones derived from the rating system. Based on this comparison, we calculated two measures:

1. Class-specific accuracy
2. Multi-class F-measure

We first calculated the accuracy for positive and negative sentiment, i.e., the proportion of positive and negative reviews rated correctly to all positive and negative reviews respectively. The results, grouped by rating, is shown in Fig. 2. At a value of between 50 - 80%, this result indicated that there was potential in our approach. Given that the classes are only three, it can be argued that a tool that randomly assigns one of the three classes can achieve up to 33.33% accuracy. For this purpose, we include a baseline for a better interpretation of the result. Also, accuracy calculated in this manner is not ideal and does not provide a reliable result since each calculation only factors in true positives and true negatives per class.

In a second experiment, we calculated the multi-class F-measure (weighing precision and recall equally), with equal weighting for precision and recall. Thus, recall identified the proportion of neutral, positive and negative reviews correctly identified against respectively all the neutral, positive and negative reviews, whereas precision identified the proportion of correctly classified (neutral, positive, negative) reviews against all reviews.

[4] www.iran-booking.com

[5] At the time of submission, Alexa lists the website as only the 7,063rd most popular in the country: http://www.alexa.com/siteinfo/www.iran-booking.com

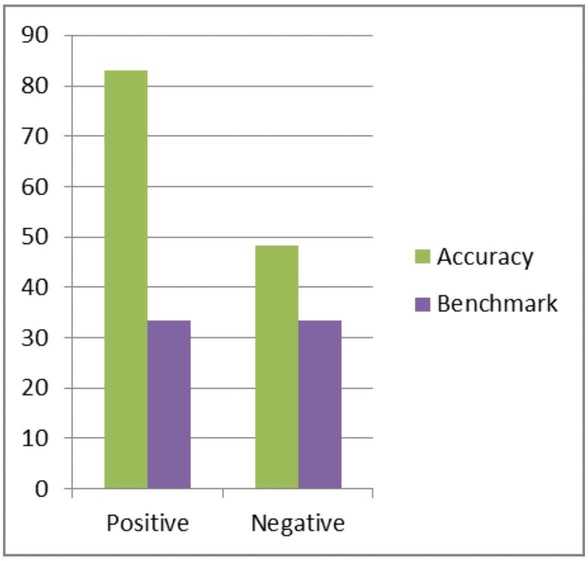

Figure 2: Overall accuracy for each rating

The resulting confusion matrix contained comparisons for the three classes and precision and recall was computed for each. The result of the three f-measures is shown in Fig. 3, again compared to the baseline. In this result, we note that although the top-performing class (positive) has gone down to just under 70%, the other two classes are not far from the 60% mark. Averaging the f-measures for the two most important classes (positive and negative), yields an average score of 68.5%.

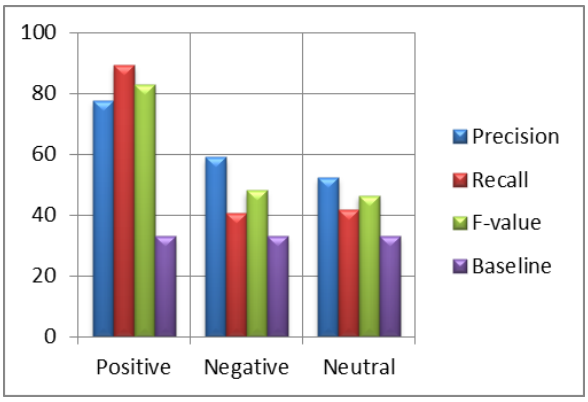

Figure 3: Multi-class F-measure

4.2 User-based Evaluation

Due to the limitations discussed above, we performed a second evaluation. In this experiment, we considered around 5100 news items from the four most popular Persian news portals (www.farsnews.com, www.tabnak.ir, www.yjc.ir www.varzesh3.com). The news items were obtained from different categories, including sport, social, politics, economic and international. For the user-based evaluation, we randomly retrieved 1170 of these items and copied them on to our website[6]. In a similar effort to the sentiment annotation phase, we circulated a request for volunteers to rate each news item. Although for the same reason as explained earlier, an exact count of volunteers is not available, website visitor IP tracking during the two weeks when the experiment was run suggests that a total of between 35-50 people have participated. This is also consistent with the appeal to rate at least 20 news items. The exercise resulted in 1116 votes for a total of 897 distinct news items. Once again, conflicting results for items with more than one vote were either resolved upon discussion (majority rule) or set to neutral. The results of manual user rating were then compared to the automatic ratings. In this case, we only focused on accuracy, starting with the user-based evaluation as the authoritative score. The results, shown in Fig. 4, show the following accuracy levels:

- positives: 67%,
- negatives: 61.8%
- neutrals: 52.5%
- overall accuracy: 60.4%
- overall accuracy (exc. neutrals): 64.4%

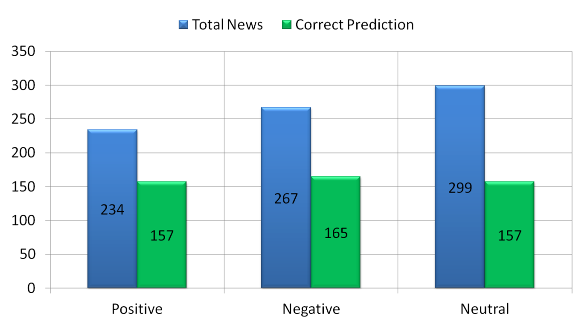

Figure 4: Performance in User-based Evaluation

5 Conclusions and Future Work

The presented approach is unique for the Persian language, since it relies on a list of entries (lexicon) paired with sentiment scores that was generated by a large number of native speakers. The approach addresses subjectivity by marking entries

[6]http://www.computerssl.com/sentiment/news.php

with conflicting scores and attempting to manually resolve said conflicts. Our experiments yield between 60-69% accuracy rates for the initial version of the lexicon-based Persian Sentiment Analysis API. Although it is still not as precise as the ML-based approach described in (Basiri et al., 2014), this compares fairly well with related work and the experiments confirm that there is value in our approach. In particular, an acceptably accurate lexicon-based approach can be used to bootstrap an ML-based system that does not require a large training set to start achieving results. Alternatively, the gazetteer could also be semi-automatically enhanced through the correction of incorrectly rated entries in a process involving human supervision. The combination of our lexicon-based approach with the most promising Persian-language ML approach to achieve a hybrid system is therefore one of the top priorities for future work. A Persian sentiment analysis API that can effectively avoid the cold-start problem when applied to a new domain can be of great value to future business use-cases. Sentiment analysis is still a highly-challenging requirement at the core of many attempts to gauge people's response or opinion about a service or product, with many use-cases in the stock market, marketing and customer care domains, as well as online customer advice. By addressing the lack of diversity in Persian sentiment analysis approaches, we want to contribute to the advancement of techniques bound to a language which remains the working language of a relatively large population. As in other languages, written Persian also faces high ambiguity in terms of context and polarity, with a high complexity also arising from mixed use of formal and informal text. In the presented research we have tried to cover both formal and informal cases in our lexicon. The evaluation indicates that there is value in our language-specific lexicon driven approach. However, a lot more remains to be done to outperform ML-based techniques and rival the list-translation (English to Persian) approach introduced by Basiri et. al. Primarily, we intend to encourage more native speakers to add and rate adjectives and phrases for the construction of a more flexible and comprehensive lexicon. In addition we also intend to improve the grammar rules to cover more of the exceptions and characteristics of the Persian language. In particular, we want to address rules centered around notorious Persian conjunctions, such as 'but and 'although. Last but not least, we also want to address abbreviated forms of writing, which is also rather common-place and which has not been addressed by the literature so far.

Acknowledgements

We would like to thank Majid Sazvar for providing us with the three own-developed linguistic components incorporated in our pipeline.

References

Muhammad Abdul-Mageed, Mona T Diab, and Mohammed Korayem. 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*, pages 587–591. Association for Computational Linguistics.

Stefano Baccianella, Andrea Esuli, and Fabrizio Sebastiani. 2010. Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In *LREC*, volume 10, pages 2200–2204.

Mohammad Ehsan Basiri, Ahmad Reza Naghsh-Nilchi, and Nasser Ghassem-Aghaee. 2014. A framework for sentiment analysis in persian.

Margaret M Bradley and Peter J Lang. 1999. Affective norms for english words (anew): Instruction manual and affective ratings. Technical report, Technical Report C-1, The Center for Research in Psychophysiology, University of Florida.

Hamish Cunningham, Diana Maynard, and Valentin Tablan. 1999. Jape: a java annotation patterns engine.

Hamish Cunningham, Diana Maynard, Kalina Bontcheva, and Valentin Tablan. 2002. A framework and graphical development environment for robust nlp tools and applications. In *ACL*, pages 168–175.

Xiaowen Ding, Bing Liu, and Philip S Yu. 2008. A holistic lexicon-based approach to opinion mining. In *Proceedings of the 2008 International Conference on Web Search and Data Mining*, pages 231–240. ACM.

Ronen Feldman, Benjamin Rosenfeld, Roy Bar-Haim, and Moshe Fresko. 2011. The stock sonarsentiment analysis of stocks based on a hybrid approach. In *Twenty-Third IAAI Conference*.

Ronen Feldman. 2013. Techniques and applications for sentiment analysis. *Communications of the ACM*, 56(4):82–89.

Hatem Ghorbel and David Jacot. 2011. Sentiment analysis of french movie reviews. In *Advances in Distributed Agent-Based Retrieval Tools*, pages 97–108. Springer.

Mohammad Sadegh Hajmohammadi and Roliana Ibrahim. 2013. A svm-based method for sentiment analysis in persian language. In *2012 International Conference on Graphic and Image Processing*, pages 876838–876838. International Society for Optics and Photonics.

Vasileios Hatzivassiloglou and Kathleen R McKeown. 1997. Predicting the semantic orientation of adjectives. In *Proceedings of the 35th annual meeting of the association for computational linguistics and eighth conference of the european chapter of the association for computational linguistics*, pages 174–181. Association for Computational Linguistics.

Yohan Jo and Alice H Oh. 2011. Aspect and sentiment unification model for online review analysis. In *Proceedings of the fourth ACM international conference on Web search and data mining*, pages 815–824. ACM.

Bing Liu. 2010. Sentiment analysis: A multi-faceted problem. *IEEE Intelligent Systems*, 25(3):76–80.

Bing Liu. 2012. Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies*, 5(1):1–167.

Andrius Mudinas, Dell Zhang, and Mark Levene. 2012. Combining lexicon and learning based approaches for concept-level sentiment analysis. In *Proceedings of the First International Workshop on Issues of Sentiment Discovery and Opinion Mining*, page 5. ACM.

Alexander Pak and Patrick Paroubek. 2010. Twitter as a corpus for sentiment analysis and opinion mining. In *LREC*, volume 10, pages 1320–1326.

Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. 2002. Thumbs up?: sentiment classification using machine learning techniques. In *Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10*, pages 79–86. Association for Computational Linguistics.

Mohamad Saraee and Ayoub Bagheri. 2013. Feature selection methods in persian sentiment analysis. In *Natural Language Processing and Information Systems*, pages 303–308. Springer.

Carlo Strapparava, Alessandro Valitutti, et al. 2004. Wordnet affect: an affective extension of wordnet. In *LREC*, volume 4, pages 1083–1086.

Afraz Z Syed, Muhammad Aslam, and Ana Maria Martinez-Enriquez. 2010. Lexicon based sentiment analysis of urdu text using sentiunits. In *Advances in Artificial Intelligence*, pages 32–43. Springer.

Maite Taboada, Julian Brooke, Milan Tofiloski, Kimberly Voll, and Manfred Stede. 2011. Lexicon-based methods for sentiment analysis. *Computational linguistics*, 37(2):267–307.

Mike Thelwall, Kevan Buckley, Georgios Paltoglou, Di Cai, and Arvid Kappas. 2010. Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, 61(12):2544–2558.

Peter D Turney. 2002. Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the 40th annual meeting on association for computational linguistics*, pages 417–424. Association for Computational Linguistics.

Taras Zagibalov and John Carroll. 2008. Automatic seed word selection for unsupervised sentiment classification of chinese text. In *Proceedings of the 22nd International Conference on Computational Linguistics-Volume 1*, pages 1073–1080. Association for Computational Linguistics.