Sentiment Analysis in Turkish Media

Cumali Türkmenoğlu and Ahmet Cüneyd Tantuğ

Istanbul Technical Universty,
Istanbul, Turkey
{turkmenogluc,tantug}@itu.edu.tr

Abstract. In this paper, we propose a comparison of Lexicon based and Machine Learning based sentiment analysis methods on Turkish social media. We formed a lexicon based SA method using a sentimentally oriented Turkish lexicon. We form this lexicon with an English opinion lexicon (translated to Turkish), multi-words expressions, words with absence/presence suffixes and extra needed words for Turkish. We explore different pre-processing techniques considering the linguistic properties of Turkish. We apply lexicon based approach by summing up sentiment scores of terms (words, MWEs, abbreviations etc.) in the lexicon. On the other hand, we formed a baseline Machine Learning (ML) based method using different feature sets including word roots and n-grams with a bag of words representation method. We apply both approaches for binary (positive/negative) classification. To show the strength and shortcomings of these two approaches, we evaluate both of them on short (twitter dataset) and long (movie dataset) Turkish informal texts.

Keywords: Sentiment Analysis, Turkish Sentiment Lexicon, Machine Learning

1 Introduction

Sentiment analysis is one of the most useful tools for social media monitoring. Automated sentiment analysis is crucial for companies' customer services to have the capability of capturing positive and negative feedbacks at the right time. Implementing an efficient sentiment analysis tool can increase the customer satisfaction and decrease the costs. The amount of accessible information with opinion on the Web has been increasing with the contribution of forums, columns, blogs, and social media. Processing this information, extracting the subjectivity and classifying the sentiment are the main challenges of the sentiment analysis. Sarcasm and irony also have remarkable importance and interest in both psychology [7] and NLP [8] [19] [15] research area. Unfortunately it is a difficult task to identify the sarcasm in a naturally occurring text even for a human [8]. We do not consider sarcasm or irony in our work.

Extracting opinions and analyzing the polarity of these opinions are the main aims of the sentiment analysis. Various approaches have been utilized to solve these problems in the past. Most of them are on subjectivity classification and sentiment classification. Subjectivity classification is the problem of classifying any segment of document as objective or subjective. On the other hand sentiment classification is the classifying of subjective sentence or document as positive or negative [10] classes according to their sentiment orientation. Subjectivity classification helps classify the sentiment [4]. NLP and machine learning techniques are extensively used for sentiment analysis. Knowing the characteristics of a language is essential for NLP and sentiment analysis because different languages require different pre-processing techniques. For instance, Turkish requires a comprehensive pre-processing step as having a productive inflectional and derivational morphology.

Sentiment analysis approaches are mainly based on ML and lexicon. Both methods have advantages and disadvantages in terms of accuracy and human labor. Our goal is to compare these methods on two different dataset types called as twitter and movie datasets. As a lexicon based method, we build a framework similar to the systems described in Thelwall et al. [20] and Vural et al. [23]. The accuracy of lexicon based method is 75.2% for Twitter dataset and 79.0% for Movie dataset. To implement ML based sentiment analysis, we have investigated several machine learning methods as Pang et al. [13] and Eroğul [6] did. We evaluate several classifiers on two different datasets with different feature sets including word roots and n-grams with bag of words representation method. The Accuracy of ML based method is 85.0% (SVM) for Twitter dataset and 89.5% for Movie dataset.

In this section we argue the importance and motivation of the sentiment analysis, general approaches and our approaches briefly. We survey related works with their approaches, success and shortages in section 2. Datasets are described in section 3. In section 4 we propose our methods in details. Section 5 presents the performance results. Finally, the paper is concluded in section 6.

2 Related Works

Pang et al. [13] conducted a supervised machine learning classification method with different features like unigrams, bigrams, POS tags, position information and combinations of them. They consider the sentiment classification problem as a non-topic based document classification and compared different machine learning techniques based on performance. They used movie reviews from IMDb to classify them as positive or negative classes. They concluded that sentiment classifying problem is more challenging than traditional topic classification. For most of ML based sentiment analysis methods the support vector machine (SVM) classifier gives the best performance.

Lexicon based approaches try to predict the sentimental orientation of an input text using sentiment scores of words and phrases in the text [22]. Lexicon based methods necessitate dictionaries which include sentimentally oriented words and/or phrases. These dictionaries can be created manually [21], or automatically, using some seed words and algorithms to expand the list with similar words [9] [22]. Manually created dictionaries are fixed. Automatically generated dictionaries eliminate the manual efforts and allow domain adaptation. Thanks to the automatic algorithms, obtaining domain specific dictionaries are relatively easier instead of building hand-crafted dictionaries for all domains.

Turney [22] presented a simple unsupervised learning algorithm to classify reviews as recommended or not recommended according to their sentimental orientation. He used mutual information between 'excellent', 'poor' and words-phrases in reviews to create sentiment lexicon automatically. Then he classified each document as recommended if the average of the sentiment orientation of words-phrases in the document is positive. A number of lexicon based studies [24] [19] showed that adjectives are the most effective terms for determining the polarity of a document.

A more recent approach is concept-level sentiment analysis, which aims to combine linguistics, lexical and machine learning techniques to improve the accuracy of sentiment classification and polarity detection. Cambria et al. [3] proposed sentic computing which exploits common-sense computing and affective computing to improve computers emotional intelligence. Poria et al. [14] introduced a concept-level approach for sentiment analysis by augmenting sentic computing framework [3] with dependency-based rules. They were able to handle texts (sentences, clauses) which have more than one topic/entity and multi-words by their approach combining with dependency-based rules. They gained 86.2% for movie dataset and 87.0% for product reviews.

The majority of sentiment analysis works are concentrated on English. However, there exists a number of sentiment analysis studies on Turkish. Eroğul [6] handled the sentimental analysis problem as a supervised machine learning classification problem and applied different ML techniques with different features like unigrams, bigrams, POS tags and combination of them [6]. His dataset is composed of movie reviews[1] and the prediction accuracy is reported as 85.0% for the binary sentiment classification. Vural et al. [23] presented a lexicon based sentiment analysis framework using Sentistrength lexicon (translated to Turkish) [20]. They used an approach based on summing lexicon scores of sentiment oriented words in related text. They used Eroğul's movie reviews and the accuracy of their framework is reported as 76.0% for positive/negative classification task. Our lexicon based method differs from [23] and [20] in terms of using multi-word expressions and absence/presence (with/without) suffixes (a kind of negation). Our ML based method is a baseline method similar to Eroğul's for comparision of these two methods.

3 Datasets

In order to evaluate the performance of lexicon and ML based sentiment analyzers, we use two datasets exhibiting different characteristics. Our first dataset is comprised of tweets which suffer from orthographical and grammatical problems. Tweets are usually difficult to process for NLP purposes since they frequently

[1] Reviews are taken from www.beyazperde.com

contain abbreviations, missing vocals that need devocalization and ungrammatical constructs both due to the character limitation of Twitter and mobile devices with limited text entry capabilities. We collect another dataset that consists of movie reviews which are more grammatical and orthographical than tweets. To show and compare the orthographical and grammatical quality of these two data sets, we present the percentage of the unique words and morphologically unrecognized words (Table 1) for each dataset. To show the quality of these datasets, we also use an editorial baseline dataset [16], which consist of news headlines. As seen in Table 1 the percentage of unique and unrecognized words is high in twitter dataset and too low in news dataset.In pre-processing level, after deascification and normalization steps, percentage of morphologically unrecognized words drops from 25.0% to 22.0% for twitter dataset, from 11.0% to 9.5% for movie dataset and from 5.0% to 4.7% for news dataset. It is clear from Table 1 that twitter dataset is more noisy than movie and news datasets.

3.1 Twitter Data

We download user generated tweets from Twitter about six different popular target entities from totally different domains. These targets consist of a famous person, an automobile brand, a football club, 2 home appliance brands and a GSM operator brand. We manually labeled these 5900 tweets as positive, negative, neutral and irrelevant. After filtering irrelevant instances, the dataset consists of 4324 tweets (pos-neg-neu) that have 14 words on the average (Table 1). In this work we use only 2978 positive and negative instances for binary classification.

Table 1. Properties of Datasets

Dataset	**Pos.**	**Neg.**	**Neu.**	**Total**	**Avg. Number of Words**	**Ratio of unrecognized words**	**Ratio of unique words in 50000 tokens**
Twitter Dataset	1677	1301	1346	4324	14	25.0%	30.0%
Movie Dataset	13224	7020	N/A	20244	38	11%	28%
News Dataset	N/A	N/A	N/A	101346	9	6.0%	23.0%

3.2 Movie Data

We create movie dataset which have different characteristics from twitter dataset to show the performance of our two different approaches. In terms of orthographical and grammatical quality, movie dataset is better than twitter dataset. We collect movie reviews from a popular website ("beyazperde.com") which presents a large number of movies from almost all film industries around the world and allows users to compose their comment about movies. Users have to select a ranking degree from 1 to 5 stars scale according to their acclaim to the target movie. Since all ranking scores do not indicate the actual sentiment of the movie review, we label them as positive when ranking score is higher than 4 and as negative when ranking score is lower than 2.5. The rest of movie reviews are rejected for reliability of the labelling process. After filtering process the number of movie reviews dropped from 60000 to 20244. Average number of words per movie review is 39 (Table 1).

3.3 Preprocessing Steps

As mentioned earlier, a number of preprocessing steps are required for both lexicon based and ML based approaches due to the productive Turkish morphology. In this study, we employ (1) deasciification and text normalization, (2) morphological analysis, (3) morphological disambiguation and (4) multi-words expressions processing.

(1) Turkish has eight Turkish specific characters which are 'ç', 'ş', 'ğ', 'ı', 'ö', 'ü' . Mobile devices with limited input capabilities cause these letters be written as standard ASCII letters as 'c', 's', 'g', 'i', 'o', 'u'. As a result, there are many tweets that are written with non-Turkish characters. These types of ASCII characters should be deasciified to their Turkish equivalents, e.g., "oluyorlarmis" should be deasciified as "ölüyorlarmış". We use Zemberek Tool [1] to handle deasciifiying operation. All positive emoticons are turned into "<]" and all negative emoticons are turned into ">[" before deasciifying step. Also additional preprocessing steps are taken for recovering on purpose misspellings like "çoook", "seviyorummm".

(2) A finite-state-machine based morphological analyzer [11] is used to produce root words, suffixes and morphological tags. This level produce ambiguous results.

(3) Since the morphological analysis stage produces ambiguous results, a morphological disambiguation module is required. We used a perceptron-based morphological disambiguation tool developed by [17].

(4) MWEs extraction aims to identify the segments of the texts which are generally sequential but not compositional [12]. We use Kemal Oflazers MWEs extraction applications Perl script to handle the MWEs extraction problem. Finally we identify and combine expressions which have different meanings and may have/havent sentiment when they separate from each other, e.g. "kafayı ye-" (literally eat the head) none of the words have an sentiment polarity by their self but it means "to get mentally deranged" and has negative sentiment polarity when they are together, Table 2.

Table 2. Some examples for Multi-words expressions

Multi-Words	**Literally**	**Meaning in English**	**Sentiment score.**
kafayı ye(mek)	(to) eat the head	get mentally deranged	-2
adam ol(mak)	(to) be a man	(to) be a good man	+2
kafayı çek(mek)	(to) pull the head	consuming alcohol	-3
ipe sapa gelmez	(he/she) does not come to rope and handle	nonsensical	-2

4 Methods

4.1 Lexicon Based Sentiment Analysis Framework

Lexicon based sentiment classification depends on comparing features of a given text with a pre-determined sentimentally oriented lexicon. Sentiment analysis does not require a detailed pre-processing [2] phase before classification for English but it is necessary for Turkish and similar agglutinating languages.

Turkish is an agglutinating language in which it is possible to add many suffixes to word roots. These derivational and inflectional suffixes can change the POS tag and sentimentally orientation of the word. An example suffixation process is presented in Table 3. Important suffixes for sentiment analysis are considered to be the negation suffix (**+ma/+me**) and absence/presence suffixes (**+sız/+siz (without), +lı/+li (with)**) which can change the sentiment orientation of a nominal word. Handling these suffixes increase the performance of the sentiment analysis [5] [13]. Text normalization pre-processing steps such as spelling correction are necessary prior to morphological analysis step since the data is noisy. The morphological analysis is needed to handle linguistic features for sentiment analysis, e.g. roots, POS tags, suffixes and adjuncts of the words. The pre-processing tasks of our framework is described in Table 4 .

Table 3. The characteristics of agglutinating Languages and Negation Suffixes

Wordform	**English Meaning**	**POS Tag**	**Sentiment Score**
iyi	good	Adj	+2
iyileş(mek)[2]	(to) improve	Verb	+2
iyileştir(mek)	(to) make sb/sth improve	Verb	+2
iyileştir**me**(mek)	not to make sb/sth improve	Verb	-2
iyileştir**me**yen	the one which does not make sb/sth improved	Adj	-2

[1] +mek is the infinitive suffix in Turkish

For a lexicon based sentiment analyzer, it is necessary to have a sentimentally oriented lexicon which is effective to detect the sentiment of a sentence. Since there were no Turkish lexicon we manually translated a basic English lexicon (Sentistrength, 2547 words) [20] into Turkish. Although there were some other more detailed lexicon in literature, such as SenticNet [3], WordNet-Affect [18], we used Sentistrength lexicon as a baseline lexicon. We reconstructed it by adding 700 MWEs, 650 words with absence/presence suffixes, 110 extra needed words for Turkish (slangs, curses and some special words) and we removed 350 root words due to adding them again as words with absence/presence suffixes. Actually we use Sentistrength as a starting point. After reconstructing, our final lexicon contains 2784 nominals and 873 verbs totally 3657 terms which have a polarity magnitude between [-5, +5].

Because of negation (-me, -ma) and absence/presence suffixes (**+sız/+siz (without), +lı/+li (with)**) suffixes, we should be careful when finding root of the words. It is not effective technique to use regular expressions like 'isolat*' which stands for 'isolate' 'isolated' 'isolation' 'isolating' in English, because of differentiation of words with suffixes in Turkish. Negation occurs in two different ways for Turkish. The first is using negation words ("değil", "yok") and second is using negation suffixes (me, -ma).

When negation suffixes met we add negation word ("değil") after related words, so that all negation forms become standardized. During calculating the sentiment score of texts, negation words change the sign of the sentiment score of the related word.

We use a booster words list ("ok", "baya", "en" etc.) which have a boosting effect when met before an adjective. We handle punctuations like '!' after sentimental terms as boosters but giving less strength.

Instead of with/without words in English we have absence/presence suffixes (**+sız/+siz (without), +lı/+li (with)**) in Turkish which are added to nouns and change their POS tag to adjective. It is a kind of negation and changes the polarity of the following word. If any absence/presence suffixes met we do not eliminate these suffixes ("umut-suz")(Table 4). As we mentioned before we also add these sentimental adjectives with absence/presence suffixes to the lexicon. So in sentiment score calculating process we compare these words with Lexicon.

4.2 Machine Learning Based Sentiment Analysis Framework

The ML approach treats the sentiment analysis as a supervised classification problem. Supervised classification requires a sufficently large labeled dataset for proper training but Lexicon based sentiment analysis does not. Determining of feature set is another key process for ML classification. In order to create the feature vector, we use unigrams and bigrams by using inverse-document-frequency (TF-IDF) feature ranking and selection method. We conduct our experiments using SVM, NB and Decision Trees (J48) classification algorithms. 10 fold cross validation technique is utilized to train and test our supervised classifiers.

Table 4. A sample input text and effects of each level of the Lexicon Based method

Methods	**A Sample Input**
Text	"Galatasaray son macini kazanamadi, maglup oldu ama umutsuz degiliz. Sevgimiz buyuk, sampiyon cimbom :) "
Deasciifying	"galatasaray son maçını kazanamadı, mağlup oldu ama umutsuz değiliz. Sevgimiz büyük, şampiyon cimbom :)"
Morphological Analysis + Morphological Disambiguation	. . . maçını maç+Noun+A3sg+P3sg+Acc kazanamadı kazan+Verb^DB+Verb+Able+Neg+Past+A3sg malup mağlup+Adj oldu ol+Verb+Pos+Past+A3sg. . . .
Multi-Words Extraction	"galatasaray son maç kazan+eylem **mağlup_ol**+eylem ama umutsuz değil sevgi büyük şampiyon cimbom >]"
Negation Handling	"galatasaray son maç **kazan+eylem değil** mağlup_ol+eylem ama umut**suz** değil sevgi büyük şampiyon cimbom >]"
Polarity detection	"galatasaray son maç kazan+eylem[2][Neg] değil mağlup_ol+eylem[-2] ama umutsuz[-3] [Neg] değil sevgi[3] büyük şampiyon[2] cimbom >][2]"
Sentiment	[+10 , -4] -> +6 (positive)

5 Results

We use accuracy measure, the number of instances that predicted correctly, to evaluate performance of our systems. We activate and deactivate modules to show the contribution of each module to performance of sentiment analyzers. Then we report results of all cases in Table 5 for Lexicon Method and Table 6 for ML Method.

According to the Table 5, each module has a contribution to the performance of Lexicon based sentiment analysis method but the most effective ones are Negation handling and MWEs handling for twitter dataset and deasciification and negation handling for movie dataset. The performance of Lexicon based sentiment analysis Method is 75.2% for twitter dataset and 79.0% for movie dataset. Results show that MWEs extraction and handling absence/presence suffixes bring reasonable improvement to performance of Lexicon based method. Since movie reviews are too long and have too many sentimental words, MWEs extraction option does not bring enough improvement.

As most researchers [10] [6] reported, our results also show that SVM has highest accuracy than other algorithms for ML approach. The best performance of ML Based Sentiment Analysis Method is 85.0% (SVM) for twitter dataset and 89.5% (SVM) for movie dataset. Using unigrams and bigrams together gives the best performance for almost all classifiers on both datasets. Results indicate that bigrams can handle most of consecutive cases such as negation, boosting and MWEs.

As surface forms of words include enough linguistic information such as negation and absence/presence suffixes, the usage of surface forms that com-

Table 5. Contribution of each module to the Performance of Lexicon Based Method

Modules	**Twitter Dataset**	**Movie Dataset**
	%	%
No ASCII converting	73.8	74.5
No disambiguation	74.5	77.0
No negation handling	72.4	76.5
No booster	74.7	77.0
No MWEs extraction	72.4	78.0
No absence/presence suffixes handling	73.7	77.0
All modules on	**75.2**	**79.0**
Only lexicon (All linguistic modules off)	68.0	71.0

bined with unigrams increases the performance of ML based method slightly for movie dataset. But it decreases the performance for twitter dataset since twitter dataset is too noisy and feature selection threshold leaves most of bigrams below the feature selection threshold (min. 20 occurrence in movie dataset and min. 5 occurrence in twitter dataset). It decreases the performance when combined with unigrams+bigrams for movie dataset also.

In comparison of these two methods, ML based method performs better than Lexicon based method on both short (twitter dataset) and long informal texts (movie dataset). The results show that accuracy of movie dataset is better than accuracy of twitter dataset in both Lexicon based and ML based sentiment analysis methods. Although Lexicon based sentiment analysis is unsupervised, it works well when text does not include sarcasm or irony.

Table 6. Contribution of each Feature set to the Performance of ML Based Method

Modules	**Twitter Dataset**			**Movie Dataset**		
	SVM%	**NB%**	**J48%**	**SVM%**	**NB%**	**J48%**
TF-IDF (Unigrams)	84.6	83.7	**81.0**	88.2	87.0	80.0
TF-IDF (Unigrams)-Surface	83.8	82.5	80.4	88.6	88.7	81.9
TF-IDF (Unigram + Bigram)	**85.0**	**84.3**	79.0	**89.5**	**89.5**	**83.0**
TF-IDF (Unigram + Bigram)-Surface	83.7	82.3	77.4	89.0	89.0	82.4

6 Conclusion

We proposed two different sentiment analysis frameworks for Turkish social media. Our framework include various pre-processing and linguistic works in order to handle the characteristics of Turkish. These pre-processing and linguistic works brings improvement to the performance of lexicon based sentiment analysis.

MWEs extraction and handling absence/presence suffixes bring reasonable improvement to performance of Lexicon based SA. It shows that discovering such hidden information (e.g. MWEs extraction and handling absence/presence suffixes) is promising, so concept-based sentiment analysis with dependency parsing could be a future work for us.

ML based method performs better than Lexicon based method on both short (twitter dataset) and long informal texts (movie dataset). Accuracy of movie dataset is better than accuracy of twitter dataset in both Lexicon based and ML based sentiment analysis methods.

Since movie dataset is more editorial and focused on target than twitter dataset, results show that both methods perform better on movie dataset. These results are also related to some topics in twitter dataset such as "Recep Tayyip Erdoğan" and "Galatasaray", to which users against them are generally in a sarcastic manner. Handling sarcasm and irony and combining it with sentiment analysis stands as a future work.

Although, ML based sentiment analysis method gives good results, being unsupervised and domain free makes lexicon based sentiment analysis preferable in many cases according the problem and dataset.

References

1. Ahmet Afsin Akın and Mehmet Dündar Akın. Zemberek, an open source nlp framework for turkic languages. *Structure*, 10, 2007.
2. Michelle Annett and Grzegorz Kondrak. A comparison of sentiment analysis techniques: Polarizing movie blogs. In *Advances in artificial intelligence*, pages 25–35. Springer, 2008.
3. Erik Cambria and Amir Hussain. Sentic computing: Techniques, tools, and applications. *Dordrecht, Netherlands, Springer*, 2012.
4. Erik Cambria, Bjorn Schuller, Yunqing Xia, and Catherine Havasi. New avenues in opinion mining and sentiment analysis. *IEEE Intelligent Systems*, 28(2):15–21, 2013.
5. Sanjiv Das and Mike Chen. Yahoo! for amazon: Extracting market sentiment from stock message boards. In *Proceedings of the Asia Pacific finance association annual conference (APFA)*, volume 35, page 43. Bangkok, Thailand, 2001.
6. Umut Eroğul. Sentiment analysis in turkish. *Middle East Technical University, Ms Thesis, Computer Engineering*, 2009.
7. Raymond W Gibbs. On the psycholinguistics of sarcasm. *Journal of Experimental Psychology: General*, 115(1):3, 1986.
8. Roberto González-Ibánez, Smaranda Muresan, and Nina Wacholder. Identifying sarcasm in twitter: a closer look. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers-Volume 2*, pages 581–586. Association for Computational Linguistics, 2011.
9. Vasileios Hatzivassiloglou and Kathleen R McKeown. 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, 1997.

10. Bing Liu. Sentiment analysis and subjectivity. *Handbook of natural language processing*, 2:627–666, 2010.
11. Kemal Oflazer. Two-level description of turkish morphology. *Literary and linguistic computing*, 9(2):137–148, 1994.
12. Kemal Oflazer, Bilge Say, et al. Integrating morphology with multi-word expression processing in turkish. In *Proceedings of the Workshop on Multiword Expressions: Integrating Processing*, pages 64–71. Association for Computational Linguistics, 2004.
13. Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. 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, 2002.
14. Soujanya Poria, Erik Cambria, Grégoire Winterstein, and Guang-Bin Huang. Sentic patterns: Dependency-based rules for concept-level sentiment analysis. *Knowledge-Based Systems, Special Issue on Big Data for Social Analysis*, 2014.
15. Antonio Reyes, Paolo Rosso, and Tony Veale. A multidimensional approach for detecting irony in twitter. *Language resources and evaluation*, 47(1):239–268, 2013.
16. Haşim Sak, Tunga Güngör, and Murat Saraçlar. Turkish language resources: Morphological parser, morphological disambiguator and web corpus. In *Advances in natural language processing*, pages 417–427. Springer, 2008.
17. Haşim Sak, Tunga Güngör, and Murat Saraçlar. Morphological Disambiguation of Turkish Text with Perceptron Algorithm, CICLing 2007, LNCS 4394, 107–118, (2007).
18. Carlo Strapparava and Alessandro Valitutti. Wordnet affect: an affective extension of wordnet. In *LREC*, volume 4, pages 1083–1086, 2004.
19. Maite Taboada, Julian Brooke, Milan Tofiloski, Kimberly Voll, and Manfred Stede. Lexicon-based methods for sentiment analysis. *Computational linguistics*, 37(2):267–307, 2011.
20. Mike Thelwall, Kevan Buckley, Georgios Paltoglou, Di Cai, and Arvid Kappas. Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, 61(12):2544–2558, 2010.
21. Richard M Tong. An operational system for detecting and tracking opinions in on-line discussion. In *Working Notes of the ACM SIGIR 2001 Workshop on Operational Text Classification*, volume 1, page 6, 2001.
22. Peter D Turney. 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, 2002.
23. A Gural Vural, B Barla Cambazoglu, Pinar Senkul, and Z Ozge Tokgoz. A framework for sentiment analysis in turkish: Application to polarity detection of movie reviews in turkish. In *Computer and Information Sciences III*, pages 437–445. Springer, 2013.
24. Janyce Wiebe. Learning subjective adjectives from corpora. In *AAAI/IAAI*, pages 735–740, 2000.