A Hybrid Approach to Generating Adjective Polarity Lexicon and its Application to Turkish Sentiment Analysis

Rahim Dehkharghani
Faculty of Engineering, University of Bonab, Bonab, Iran
Email: rdehkharghani@bonabu.ac.ir

Abstract—Many approaches to sentiment analysis benefit from polarity lexicons. Existing methods proposed for building such lexicons can be grouped into two categories: (1) Lexicon based approaches which use lexicons such as dictionaries and WordNet, and (2) Corpus based approaches which use a large corpus to extract semantic relations among words. Adjectives play an important role in polarity lexicons because they are better polarity estimators compared to other parts of speech. Among natural languages, Turkish, similar to other non-English languages suffers from the shortage of polarity resources. In this work, a hybrid approach is proposed for building adjective polarity lexicon, which is experimented on Turkish combines both lexicon based and corpus based methods. The obtained classification accuracies in classifying adjectives as positive, negative, or neutral, range from 71% to 91%.

Index Terms—Sentiment analysis, Polarity Lexicons, Adjectives.

I. INTRODUCTION

Due to ever-increasing amount of online information in social media, manual processing of data to extract valuable information is impractical. Information retrieval automatically extracts the desired information from (usually large) datasets. This extracted information might be the embedded sentiment polarity in data (generally texts). The work of sentiment analysis is thus to extract the sentiment polarity (positivity and negativity) from data.

Many approaches to sentiment analysis require polarity lexicons to assign a polarity tag (positive, negative or neutral) to a segment of text. There exist a good deal of work on polarity lexicon generation which is grouped into two categories by Liu [1]: lexicon based methods and corpus based methods. Lexicon based methods start with a small seed word list and expand it upon synonymy and antonymy relations using dictionaries such as WordNet [2]. In corpus based methods, semantic relations between terms in a corpus are employed to generate polar terms. These relations include pointwise mutual information [3],

co-occurrence of words, conjoined adjectives (by "and" or "but") [4], and delta tf-idf [5].

The above mentioned methods have been separately used in the literature; however, they could be combined to design a more effective approach which is accomplished in this work. Each method contributes to the proposed hybrid approach as a classification feature in classifying adjectives into positive, negative, or neutral. Experimental evaluation approves the effectiveness of the hybrid approach compared to each method in isolation. The proposed method in this paper can be applied on other languages with some changes.

II. TURKISH AND ITS CHALLENGES FOR SENTIMENT ANALYSIS

Turkish is a member of the Turkic family of Altaic languages. Particular characteristics of Turkish make natural language processing (NLP) and sentiment analysis tasks difficult for this language. Morphologically, Turkish is an agglutinative language with morphemes attaching to a root-word as "beads-on-a-string". Words are formed by very productive affixations of multiple suffixes to root words, from a lexicon of about 30K root words (not counting proper names.) Nouns do not have classes nor are there any markings of grammatical gender in morphology and syntax. When used in the context of a sentence, Turkish words can take many inflectional and derivational suffixes. It is quite common to construct words which correspond to almost a sentence in English: For example, the equivalent of the Turkish word "sağlamlaştırabileceksek" in English can be expressed with the fragment if we will be able to make [it] become strong (fortify it) [6].

For Turkish, the morphological structure of a word is also necessary for sentiment analysis in addition to the root word, as suffixes may change the polarity of a word. For instance, the word iştahsız (having no appetite), is negative (due to suffix +sız), while its antonym, iştahlı, is positive (due to suffix +lı). Note that the root word itself, istah (appetite), is mostly positive.

In terms of NLP tools for Turkish, there are morphological analysers [7], and parsers [8].

Morphological analysers perform a word-level analysis and provide the root word but parsers provide more information such as the dependency parse tree of a sentence and the POS tag of words according to their context in the sentence.

III. RELATED WORK

Sentiment analysis can be considered in different levels of granularity: document, sentence, phrase, and word levels. In [9], the author investigated document level sentiment analysis using machine learning techniques. In sentence level [10], Meena and Prabhakar addressed the effect of conjunctions, and semantic relations between sentences. In phrase level sentiment analysis, two works have been accomplished by Wilson and colleagues [11, 12]. In [11], the authors proposed an approach which first classifies an expression as subjective or objective and then estimates its polarity in the case of subjectivity. This method estimates the contextual polarity of an expression by using a large number of subjectivity clues and the prior polarity of appeared words in the expression. This work mostly relies on statistical methods. The authors extended their work in [12]. The focus of this work is to figure out which features are more important in automatically distinguishing between prior and contextual polarity. Multi-perspective Question Answering (MPQA) is used as the opinion lexicon in this work. Again in phrase level, Agrawal et al. [13] proposed a method to predict contextual polarity of subjective phrases in a sentence. The authors present new classification features which could achieve higher accuracies in ternary (positive/negative/neutral) classification of phrases over two baselines–majority class baseline as well as a more difficult baseline consisting of lexical n-grams. Yi et al. [14] analyzed grammatical sentence structures and phrases for sentiment analysis purposes. The authors present Sentiment Analyzer which extracts sentiment about a subject from online text documents. Instead of classifying the sentiment of an entire document about a subject, the designed system detects all references to the given subject, and determines the sentiment in each of the references. Nasukawa and Yi [15] proposed an approach for extracting sentiments associated with positive or negative polarity for specific subjects in a document, instead of classifying the whole document as positive or negative. In this work, the goal is to identify semantic relationships between sentiment expressions and subject terms. The authors mainly focus on increasing the precision even if the recall remains low. Trend analysis and important document identification in terms of sentiments are two examples that can take advantage of this work. In concept-level, Tsai et al. [16] presented a two-step method which combines iterative regression and random walk with in-link normalization to build a concept-level sentiment lexicon. Poria et al. [17] presented a methodology for enriching SenticNet [18]–a polarity lexicon in English–concepts with affective information by assigning an emotion label to those concepts.

There exist also a good deal of research on building polarity lexicons. Liu [1] categorizes these methods into two groups: lexicon based approaches and corpus based approaches. Lexicon based approaches start with a small seed set (e.g., 20 terms) and expand the list by using the existing relations such as synonymy and antonymy among terms in dictionaries. Hu and Liu [19] used this method to generate a list of polar English terms and then manually cleaned up the generated list to remove errors. The same approach was used in [20] to build a polarity lexicon for Turkish. A similar approach was proposed in [21], which assigns also a sentiment score to each word by using a probabilistic method. Esuli and Sebastiani [22] proposed a methodology to assign three polarity scores (positive, negative, and neutral) to each synset in English WordNet. This approach was modified in [20] to build a polarity lexicon for Turkish based on the Turkish WordNet [23]. In corpus based approaches, having a seed word list with known polarities, new polar words are extracted based on the existing semantic relations in the corpus. One of the early ideas was proposed in [4]. The authors used conjunctions in a corpus to find new polar adjectives. They show that conjoined adjectives by "and" usually have the same polarity while they usually have opposite polarity when conjoined by "but". Some extra relations such as "Either or" and "Neither-nor" were also used for this purpose. This assumption holds also for Turkish as experimented in the current work. Kanayama and Nasukawa [18] followed this approach and improved it by adding this idea: "consecutive sentences usually have the same polarity". Another popular method was proposed in [3], which introduced Pointwise Mutual Information (PMI). Turney computed the PMI score of adjectives with "excellent" as a pure positive and also with "poor" as a pure negative word, co-occurred in a sequence of words as a window. Wu and Wen [24] deal with the problem of comparative sentences in Chinese by relying on the proposed method by Turney and also Web search hit counts. There exist some other work with a focus on specific issues of sentiment lexicons. For example, Bosco et al. [25] focused on Irony which is a challenging issue in developing sentiment lexicons. The authors present SentiTUT corpus–a collection of texts from Twitter annotated with sentiment polarity.

There are also some effort on sentiment analysis of Turkish text. Yıldırım et al. [26] accomplished a sentiment analysis task on Turkish tweets in the telecommunication domain. The authors applied a multi-class ternary (positive/negative/neutral) classification by support vector machines on tweets using features such as inverse document frequency, unigrams and adjectives. They also benefit from NLP techniques such as Normalization, stemming, and negation handling. Vural et al. [27] presented a system for unsupervised sentiment analysis in Turkish text documents using SentiStrength [28] by translating its polarity lexicon to Turkish. SentiStrength is a sentiment analysis tool for English which assigns a positive and a negative score to a segment of text. Kaya et al. [29] investigated sentiment analysis of Turkish political news in online media. The

authors used four different classifiers–Naive Bayes, Maximum Entropy, SVM, and the character based n-gram language models– with a variety of text features: frequency of word unigrams, bigrams, root words, adjectives and effective (polar) words. They conclude that the Maximum Entropy and the n-gram language models are more effective than the SVM and Naive Bayes classifiers in classifying Turkish political news. Boynukalın [30] has worked on emotion analysis of Turkish texts by using machine learning methods. The author has investigated four types of emotions: joy, sadness, fear, and anger on a dataset that she built for this purpose. Eroğul [31] explored the use of surface linguistic features such as part-of-speech (POS) tags, word unigrams and bigrams, and negation markers in sentiment analysis. This work relies on Zemberek [7], a morphological analyzer for Turkish.

IV. Polarity Lexicon Generation

In this work, in order to generate an adjective polarity lexicon, a list of 11,000 Turkish adjectives is downloaded from a Turkish webpage—unfortunately it is out of order at the moment. Only unigrams and bigrams are covered; trigrams adjectives are very scarce. A bigram adjective (adjective phrase) is composed of two words appearing together as an adjective, e.g., "akla yatkın" (advisable, literally means approvable by the wisdom). The proposed methodology receives a list of raw adjectives as input and classifies them as three classes (positive, negative, and neutral), while the majority of the existing approaches extract these adjectives from linguistic corpora or lexicons. Different methods have been used in adjective classification, each of which contributes to the classification tasks as a feature. These features are explained in the following section.

4.1 Classification Features for Adjectives

In this section, several polarity estimator methods are explained, which are used as features in classifying adjectives into polarity classes.

Pointwise Mutual Information (PMI): This method captures the co-occurrence frequency of two words in a corpus. The main idea is that positive words generally co-occur with positive adjectives and negative ones co-occur with negative adjectives. PMI was first proposed by Turney [3] for computing the co-occurrence frequency of words with two positive and negative words: "excellent" and "poor". He proposed equation (1) for computing PMI score of two words, which is modified in this work as (2) for computing the PMI score of one adjective with a list of positive (or negative) words. In equation (2), P(w1) is the probability of seeing w1 and P(w1,w2) is the probability of seeing both w1 and w2 in a specified window (e.g., a sentence). In equation (2), the average PMI score of each adjective separately with 1,000 positive and 1,000 negative words that were already generated for Turkish in previous work [32] is computed.

This co-occurrence is searched among 270,000 Turkish sentences in Turkish movie reviews1 as the corpus.

$$PMI(w_1, w_2) = log\left(\frac{P(w_1, w_2)}{P(w_1), P(w_2)}\right) \qquad (1)$$

$$\sum_{adj \in PWS} \frac{log(P(adj, w_i))}{p(adj)p(w_i)} \qquad (2)$$

$P(adj, w_i)$ is the probability of seeing both adj and w_i in the corpus, and similarly, $P(adj)$ and $P(w_i)$ are respectively the probability of seeing the adj and wi in the copus, in isolation.

Delta document frequency: In this technique, the normalized document frequency of an adjective in negative sentences is subtracted from its normalized document frequency in positive sentences. Equation (3) is used for computing this score. Each adjective is searched among 2,700 sentences randomly selected from Turkish movie reviews and labeled as positive, negative, or neutral. The details of labeling task is explained in Section 5.1.

$$delta - df(adj, s) = \frac{adj \in s; s \in S_{pos}}{|S_{pos}|} - \frac{adj \in s; s \in S_{neg}}{|S_{neg}|} \qquad (3)$$

adj stands for a given adjective, s for a sentence and S for a dataset of sentences. S_{pos} and S_{neg} are the set of positive and negative sentences. The intuition behind this method is that the greater the number of positive (or negative) sentences including adj, the greater the chance of positivity (or negativity) for adj.

Translating to English: In this feature, all adjectives are translated to English by a bilingual dictionary[2]; then the sentiment polarities of first three English translations of each Turkish adjective are estimated based on three polarity lexicons in English: Polar word list [20], SentiWordNet [22], and SenticNet [18]. In Polar word list, positive words have been already separated from negative ones. In SentiWordNet, a word is assumed as positive (in this work) if the average positive polarity of all synsets of the word separated by POS tags is higher than its negative score. Word Sense Disambiguation (WSD) is not investigated more deeply in this paper, since it is an ongoing problem in Turkish. In SenticNet, if the overall polarity of a word is positive (or negative), it is assumed as a positive (or negative) word. Note that the weight of the i_{th} translation is assumed two times greater than the weight of $(i + 1)_{th}$ translation in this feature. Finally, a Turkish word is labelled as positive (or negative), if English polarity lexicons label it as positive (or negative) by using the majority voting method. This feature has been used as the baseline for adjective polarity lexicon generation.

Hit number in Google: In this feature, the expressions "*adj ve güzel*" (adj and good/beautiful), and "*adj ve kötü*" (adj and bad) are searched in Google search engine, where *adj* is an adjective in the adjective list. As

---

1 This dataset is collected from www.beyazperde.com
2 www.seslisozluk.net

conjoined adjectives connected by "*ve*" (and) generally have the same polarity, an adjective is expected to be positive, if its hit number in Google for the phrase "*adj ve güzel*" is greater than that of the phrase "*adj ve kötü*", or negative otherwise. Equation (4) is used for this purpose. *hit(phrase)* gives the number of hits in Google returned for the searched *phrase*.

$$DeltaHit(adj) =$$
$$\log(hit(adj\ ve\ güzel) - hit(adej\ ve\ kötü)) \quad (4)$$

Table 1 lists the above-mentioned classification features plus linguistic techniques (conjunctions and suffixes) for classifying the adjectives.

Table 1. Classification features and linguistic techniques for classifying adjectives

Classification features
Delta Doc-frequency
Hit number in Google
Translating to English
Pointwise mutual information
Linguistic techniques
Conjunctions
Suffixes

4.2 Classification of Adjectives

In this step, suggested features in Section 4.1 are combined to train a classifier. For this purpose, 1,100 adjectives (10% of all data) are manually labelled by native speakers as positive, negative, or neutral and their feature vectors as well as their polarity labels are fed to the classifier. The labeling task is accomplished by three (plus one) Turkish native speakers; their agreement on the label of adjectives is 87%. This means that at least two labellers agreed on 87% of adjectives and the remaining 13% were judged by the fourth labeller who had to agree with one of the three labelers. The fourth labeler helped label those samples which were not agreed by at least two labelers. This policy is used in all labelling tasks of the current work. Then the polarities of adjectives in the remaining set (9,900 adjectives) are estimated by classifying them as one of the above mentioned three classes. At the end of this phase, 1,490 positive, 1,210 negative, and 7,300 neutral adjectives are obtained. Adjective classification has been accomplished by three classification tasks: ternary classification of all adjectives into positive, negative, or objective, and two binary classifications: classifying all adjectives as objective and subjective and then classifying subjective adjectives as positive and negative. The tool used for classification tool is WEKA [33] and the classifiers used for classification are three state of the art classifiers namely, Multilayer perceptron, Logistic Regression, and SVM. These classifiers have been tested on all classification tasks but finally the best one for each task has been chosen. In ternary classification of adjectives, the generalization accuracy of Logistic regression is higher than the other classifiers but in the other two (binary) classification tasks (positive/negative and objective/subjective), SVM gives slightly better results.

Afterwards, the obtained polarity lexicon is expanded by two linguistic techniques explained in the following subsection. Experimental result for evaluating the proposed classifier can be seen in Section 5.2.

4.3 Improvement Phase on Classification

This phase consists of two tasks: (1) adding new polar adjectives by using conjunctions, and (2) adding new polar adjectives by adding/removing suffixes to/from already generated adjectives.

4.3.1 Conjunctions for Adjective Extraction

As mentioned earlier, conjoined adjectives by "*ve*" (and) are expected to have the same polarity; however, they usually have opposite polarity when conjoined by "*ama*" (but). Using this method, all conjunctions are extracted from 270,000 Turkish sentences in Turkish movie reviews. The generated lexicon is then expanded by new polar adjectives based on patterns listed in Table 2. In this table, if the polarity of an adjective (*adj*) in one side of the conjunction is unknown, but the other side (*NegAdj* or *PosAdj*) is known, a polarity tag is assigned to *adj* based on the above mentioned patterns. The polarity of an adjective is known if it is a member of manually labelled training set or its polarity can be extracted from the existing Turkish polarity lexicons including SentiTurkNet [20] and Polar Word Set [20]. Therefore, if an adjective is not known, it cannot estimate the polarity of unknown adjectives. Among 270,000 Turkish sentences, 1432 conjunctions of adjectives were found and after applying the conjunction technique, 352 new adjectives were extracted. As the accuracy of conjunction technique (91.7) is higher than the classification (71.7), in the case of generating an adjective by two techniques, the assigned polarity by conjunctions is preferred. By using the conjunction technique, the polarity of 121 adjectives were updated. In summary, this technique could generate 98 new positive and 102 new negative adjectives that did not already exist in the polarity lexicon.

Table 2. Patterns used for extracting new polar adjectives by conjunctions.

Negative adjective extraction	Positive adjective extraction
adj ve *NegAdj*	*adj* ve *PosAdj*
NegAdj ve *adj*	*PosAdj* ve *adj*
adj ama *PosAdj*	*adj* ama *NegAdj*
PosAdj ama *adj*	*NegAdj* ama *adj*

4.3.2 Suffixes for Adjective Extraction

In Turkish, as an agglutinative language, suffixes can be added to the root word to build new words, such as adding suffixes to nouns to generate adjectives. A suffix can also change the polarity of a word; the Turkish noun "*kullanıs*" (usage), for example, is neutral; its polarity changes to negative by adding the suffix "*sız*" (-less): "*kullanıssız*" (useless), but remains positive due to the suffix "*lı*" (-ful) in "*kullanıslı*" (useful). One method for assigning a polarity tag to a Turkish word is to decompose it into the root word and suffixes. In this

method, the root word is searched in polarity lexicons and the polarity of the word changes if adding suffixes shifts its polarity. For example, "*sevgisiz*" (loveless) has the root "*sev*" (love) as infinitive verb form, and the suffix "*gi*" transforms the verb to noun with the same meaning and polarity but the suffix "*siz*" (without) transforms the noun to an adjective and flips the polarity. Another approach to extract the polarity of this kind of words is to add the whole word (e.g., *sevgisiz*) to a polarity lexicon. The second method is followed in this work by generating a large number of polar words to be used in Turkish sentiment analysis systems. In this work, only two sets of suffixes are used which are listed in Table 3, because they generate new polar words with a negligible error rate. For example, if an adjective ended by (li, lı, lu, l ü) is positive, replacing the suffix with [*siz, sız, suz, süz*] (-less) will generate a negative adjective. Note that there are many other suffixes for transforming nouns and adjectives to each other, but almost all of them generate erroneous (or irrelevant) words when adding/removing them to/from words. As an erroneous example, the suffix "*sel*" (related to) can transform a noun to an adjective, but in this work, if it is used to obtain new adjectives from nouns, the newly generated words must be manually reviewed to check their validity in Turkish. Therefore, only the mentioned suffixes in Table 3 are used, which could achieve the highest accuracy in automatic generation of new polar adjectives by replacing a suffix with another. By employing this method, 249 positive and 151 negative adjectives were generated but only 205 of them were not already generated.

Furthermore, a small set of polar nouns (108 words) is also generated by removing suffixes from adjectives. For example, by removing the suffix "*sız*", the negative adjective "*heyecansız*" (without excitement) changes to positive noun "*heyecan*" (excitement). The small polar noun set generated by this technique can also be useful for sentiment analysis systems.

Table 3. Patterns used for extracting new polar adjectives by changing their suffixes.

Adj+(lı)=> Adj+(sız) [pos=>neg]: insaflı => insafsız
Adj+(lu) => Adj+(suz) [neg=>pos]: korkulu => korkusuz
Adj+(siz) => Adj+(li) [pos=>neg]: kedersiz => kederli
Adj+(sız) => Adj+(lı) [neg=>pos]: vicdansız => vicdanlı

V. Experimental Evaluation

This section evaluates the proposed methodology by classification accuracy, confusion matrix (per-class accuracy), and an extrinsic evaluation. As there exist two publicly available Turkish polarity lexicons: polar word set, and SentiTurkNet, it might be interesting to compute intersection of the generated lists with the existing ones. Actually, 31% of the words in newly generated adjective polarity lexicon is in common with the existing polarity lexicons but the remaining adjectives (69%) are new polar words.

5.1 Dataset

As mentioned in previous sections, the proposed approach has been applied on documents of movie domain in Turkish, which are more formal than some other data types such as microblogging data but it could be applied on other domains such as restaurant, hotel, and communication reviews; however, applying it on other kinds of textual data such as social media would be challenging, as those data (e.g., tweets) are usually informal and noisy, including abbreviations and useless text.

Sentiment analysis is a domain-dependent task, therefore a given term may have different polarities in different domains; e.g., the term "big" is positive for room size in hotel domain but negative for battery size in camera domain.

In this paper, the list of adjectives are domain-independent. Although extracting polar terms from one domain and applying on another may have some drawbacks, however, in resource-lean languages such as Turkish we have to accept these weaknesses in polarity lexicons because the lack of resources (especially domain-dependent documents) is a big problem in these languages. In this work, two datasets have been used.

Turkish movie reviews[3]. We have manually labelled 1,000 randomly chosen documents from this dataset as positive, negative, or neutral in our previous work [32]. We also labelled 2,700 sentences appearing in these documents as positive, negative, or neutral. Only the labels of sentences are used in this work. The distribution of [neutral, positive, and negative] sentences and documents are [50%, 30%, 20%] and [52%, 29%, 19%] respectively. The average length of each document and each sentence in this domain are respectively 23 and 9 words. The labelling task is accomplished by three (plus one) people and the agreement among at least two labellers is 84% for sentence level analysis. Again, the fourth labeller helped tag those sentences which were not agreed by at least two labellers.

Turkish restaurant reviews. This dataset was used as training set in Semeval 2016-task 5 [34], which has been already labelled with three tags: positive, negative, and neutral. The aim of task 5 in Semeval competition is to estimate the polarity label of each aspect appearing in a sentence. This dataset includes 239 documents and 1104 sentences, which has been used in the current work for evaluating the usefulness of the generated lexicon in sentence-level sentiment analysis. The average length of each document and each sentence in this domain are respectively 26 and 8 words.

5.2 Experiments

Evaluation of the proposed approach includes investigating the effect of each feature in each classification task which is presented in Table 4. Confusion matrices for both binary and ternary

---

[3] These reviews are collected from www.beyazperde.com which are available in http://sentilab.sabanciuniv.edu/wpcontent/uploads/2015/03/TurkishMovieReviews.txt

classification tasks are provided in Tables 5, 6, and 7. "Direct translation to English" (translation feature) is supposed as a baseline method and the proposed methodology outperforms the baseline by about 17 percentage points in classification accuracy. The classification method is Logistic classifier and evaluation is done by 5-fold cross-validation. In this evaluation method, data is divided into five equal parts; 80% of data is used to train the classifier, and the remaining 20% is used to test the classification system. Tjis task is repeated five times with different 80/20 percent odf data and the average accuracy of these tasks is used as the overall accuracy of the classifier.

Table 4. Binary and ternary classification accuracies for adjectives by Logistic Classifier (ternary classification) and SVM (binary classification) using 10-fold Cross Validation on training set(%)

Feature-name	Ternary	obj/subj	pos/neg
Hit number	52.45	53.67	53.03
Delta tf-idf	54.24	57.93	60.02
Translations	54.20	56.70	56.87
PMI	53.20	56.75	56.67
All features	71.77	72.61	73.95
Conjunctions	91.70	-	-
Suffixes	73.50	-	-

Table 5. Confusion matrix for binary (positive and negative) classification of adjectives with all features.

		Estimated	
		Positive	Negative
True	Positive	0.79	0.21
	Negative	0.30	0.70

Table 6. Confusion matrix for binary (subjective and objective) classification of adjectives with all features.

		Estimated	
		subjective	objective
True	Subjective	0.80	0.20
	Objective	0.25	0.75

Table 7. Confusion matrix for ternary (positive, negative, and neutral) classification of adjectives with all features.

		Estimated		
		Positive	negative	Neutral
True	Positive	0.74	0.10	0.16
	Negative	0.13	0.67	0.20
	neutral	0.10	0.18	0.72

5.3 Extrinsic evaluation

In order to evaluate the usefulness of the generated polarity lexicon, an extrinsic evaluation is carried out. The generated lexicon as well as other two Turkish lexicons, polar word set and SentiTurkNet, are used to estimate the polarity of Turkish restaurant reviews. This dataset includes 1104 Turkish sentences in restaurant domain. This dataset has been used as a benchmark in Semeval competition task5-Aspect based sentiment analysis. In this task, the goal is to estimate the polarity of aspects appearing in a sentence which have been tagged with three labels: positive, negative, and neutral. The obtained accuracies with and without using the

generated polarity lexicon are given in Table 8. In this table, the abbreviations STN, PWS, and AL respectively stand for SentiTurkNet, Polar Word Set, and Adjective Lexicon. This sentiment analysis task simply searches for polar words in a sentence. No NLP technique except tokenization and word cleaning is employed in this system, as the goal is only to evaluate the usefulness of generated lexicon. As seen in Table 8, adding the adjective polarity lexicon increases the classification accuracy by three percentage points. Note that for generating adjective lexicon, no dataset except movie reviews have been used.

Table 8. The accuracy of Binary and ternary classification of Turkish restaurant reviews by Logistic Classifier (ternary classification) and SVM (binary classification) using 10-fold Cross Validation on training set(%).

Lexicons	Binary	Ternary
STN +PWS	73.22	67.23
STN+PWS+AL	76.17	70.81

lexicons using one or some of the features (used in this work) in classification of adjectives. Turney [3], who proposed the PMI concept, achieved accuracies ranging from 66% in Movie reviews to 84% in automobile reviews in classifying the reviews as positive and negative. Hatzivassiloglou and McKeown [4] proposed the idea of conjoined adjectives by "*and*" and "*but*" which resulted in accuracies ranging from 78% to 82% in classifying adjectives as positive and negative extracted from 21 million word 1987 Wall Street Journal corpus [35]. Note that the above mentioned works have been applied on English while the current work is the first one applied on Turkish. According to the results reported in Section 5, the following conclusions can be extracted.

-The proposed approach for building adjective polarity lexicon, outperforms the baseline approach–direct translation to English– by four percentage points. This issue approves that building polarity lexicons specifically for a non-English language is more efficient than translating English lexicons to non-English languages. This assumption was also approved in previous work [20].

- The highest per-class accuracies (confusion matrix values) belong to the positive class and lowest accuracies belong to the negative class. Generally positive expressions are more clearly expressed by people, compared to the negative expressions.

VI. CONCLUSION AND FUTURE WORK

In this work, a hybrid methodology is proposed to build a polarity lexicon including adjectives which are experimented on Turkish. The proposed methodology consists of several methods such as point-wise mutual information and conjoined adjectives. The mentioned techniques have been separately used in the literature, while in this work, they are combined to provide a more effective methodology. The obtained accuracies in classifying adjectives as three classes range from 71% to

91%. Linguistic techniques such as adding or removing suffixes to/from words are also applied to achieve better performance. The future work includes adding adverbs to existing polarity lexicons in Turkish. Although adjectives are proven to be good, it is also proven that adjectives plus adverbs are better than adjectives alone [36].

REFERENCES

[1]   Liu, B. Sentiment Analysis and Opinion Mining. Morgan and Claypool Publishers, USA, 2012.

[2]   Miller, G. A. WordNet: a lexical database for English. Communications of the ACM 38, 11 (1995), 39–41.

[3]   Turney, P. D. 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 (2002), Association for Computational Linguistics, pp. 417–424.

[4]   Hatzivassiloglou, V., and McKeown, K. R. Predicting the semantic orientation of adjectives. In Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics (1997), pp. 174–181.

[5]   Martineau, J., and Finin, T. Delta tfidf: An improved feature space for sentiment analysis. In ICWSM (2009).

[6]   Oflazer, K., and Bozsahin, H. C. Turkish natural language processing initiative: An overview. In Middle East Technical University (1994), Citeseer.

[7]   Akin, A. A., and Akin, M. D. Zemberek, an open source NLP framework for Turkic languages. Structure 10 (2007).

[8]   Eryigit, G. ITU Turkish NLP web service. In Proceedings of the Demonstrations at the 14th Conference of the European Chapter of the Association for Computational Linguistics (EACL) (Gothenburg, Sweden, April 2014), Association for Computational Linguistics.

[9]   Pang, B., Lee, L., and Vaithyanathan, S. Thumbs up? sentiment classification using machine learning techniques. In Proceedings of EMNLP (2002), pp. 79–86.

[10] Meena, A., and Prabhakar, T. Sentence level sentiment analysis in the presence of conjuncts using linguistic analysis. Springer Berlin Heidelberg, 2007.

[11] Wilson, T., Wiebe, J., and 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 (2005), Association for Computational Linguistics, pp. 347–354.

[12] Wilson, T., Wiebe, J., and Hoffmann, P. Recognizing contextual polarity: An exploration of features for phrase-level sentiment analysis. Computational Linguistics (2009), 399–433.

[13] Agarwal, A., Biadsy, F., and Mckeown, K. R. Contextual phrase-level polarity analysis using lexical affect scoring and syntactic n-grams. In Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics (2009), Association for Computational Linguistics, pp. 24–32.

[14] Yi, J., Nasukawa, T., Bunescu, R., and Niblack, W. Sentiment analyzer: Extracting sentiments about a given topic using natural language processing techniques. In ICDM 2003, Third IEEE International Conference on (2003), IEEE, pp. 427–434.

[15] Kanayama, H., and Nasukawa, T. Fully automatic lexicon expansion for domain-oriented sentiment analysis. In Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing (2006), Association for Computational Linguistics, pp. 355–363.

[16] Tsai, A. C.-R.,Wu, C.-E., Tsai, R. T.-H., Hsu, J. Y.-j., et al. Building a concept-level sentiment dictionary based on commonsense knowledge. IEEE Intelligent Systems 28, 2 (2013), 22–30.

[17] Poria, S., Gelbukh, A., Hussain, A., Howard, N., Das, D., and Bandyopadhyay, S. Enhanced senticnet with affective labels for concept-based opinion mining. IEEE Intelligent Systems 28, 2 (2013), 31–38.

[18] Cambria, E., Olsher, D., and Rajagopal, D. Senticnet 3: a common and common-sense knowledge base for cognition-driven sentiment analysis. In Twenty-eighth AAAI conference on artificial intelligence (2014).

[19] Hu, M., and Liu, B. Mining and summarizing customer reviews. In Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining (2004), ACM, pp. 168–177.

[20] Dehkharghani, R., Saygin, Y., Yanikoglu, B., and Oflazer, K. Sentiturknet: a Turkish polarity lexicon for sentiment analysis. Language Resources and Evaluation 53 (2015), 1–19.

[21] Kim, S.-M., and Hovy, E. Determining the sentiment of opinions. In Proceedings of the 20th international conference on Computational Linguistics (2004), Association for Computational Linguistics, p. 1367.

[22] Baccianella, S., Esuli, A., and Sebastiani, F. Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In LREC (2010), vol. 10, pp. 2200–2204.

[23] Bilgin, O., Çetinoglu, Ö., and Oflazer, K. Building a wordnet for Turkish. Romanian Journal of Information Science and Technology 7, 1-2 (2004), 163–172.

[24] Wu, Y., and Wen, M. Disambiguating dynamic sentiment ambiguous adjectives. In Proceedings of the 23rd International Conference on Computational Linguistics (2010), Association for Computational Linguistics, pp. 1191–1199.

[25] Bosco, C., Patti, V., and Bolioli, A. Developing corpora for sentiment analysis: The case of irony and senti-tut. IEEE Intelligent Systems 28, 2 (2013), 55–63.

[26] Yıldırım, E., Çetin, F. S., Eryigit, G., and Temel, T. The impact of NLP on Turkish sentiment analysis. TÜRKIYE BILISIM VAKFI BILGISAYAR BILIMLERI ve MÜHENDISLIGI DERGISI 7, 1 (Basılı 8 (2015).

[27] Vural, A. G., Cambazoglu, B. B., Senkul, P., and Tokgoz, Z. Ö. A framework for sentiment analysis in Turkish: Application to polarity detection of movie reviews in Turkish. In ISCIS (2012), E. Gelenbe and R. Lent, Eds., Springer, pp. 437–445.

[28] Thelwall, M., Buckley, K., and Paltoglou, G. Sentiment strength detection for the social web. Journal of the American Society for Information Science and Technology 63, 1 (2012), 163–173.

[29] Kaya, M., Fidan, G., and Toroslu, I. H. Sentiment analysis of Turkish political news. In Proceedings of the The 2012 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technology-Volume 01 (2012), IEEE Computer Society, pp. 174–180.

[30] Boynukalin, Z. Emotion analysis of Turkish texts by using machine learning methods. MSc thesis, Middle East Technical University (2012).

[31] Eroğul, U. Sentiment analysis in Turkish. MSc thesis, Middle East University, Turkey (2009).

[32] Dehkharghani, R., Yanikoglu, B., Saygin, Y., , and Oflazer, K. Sentiment analysis in Turkish at different granularity levels. Natural Language Engineering (2016).

[33] Holmes, G., Donkin, A., andWitten, I. H. Weka: A machine learning workbench. In Intelligent Information

Systems, 1994. Proceedings of the 1994, Second Australian and New Zealand Conference on (1994), IEEE, pp. 357–361.

[34] Pontiki, M., Galanis, D., Papageorgiou, H., Androutsopoulos, I., Manandhar, S., Al-Smadi, M., Al-Ayyoub, M., Zhao, Y., and Qin, B. Orphée de clercq, véronique hoste, marianna apidianaki, Xavier tannier, natalia loukachevitch, evgeny kotelnikov, nuria bel, salud marıa jiménez-zafra, and gülsen eryigit. 2016. semeval-2016 task 5: Aspect based sentiment analysis. In Proceedings of the 10th international Workshop on Semantic Evaluation, SemEval (2016), vol. 16.

[35] Marcus, M. P., Marcinkiewicz, M. A., and Santorini, B. Building a large annotated corpus of English: The penn treebank. Computational linguistics 19, 2 (1993), 313–330.

[36] Benamara, F., Cesarano, C., Picariello, A., Recupero, D. R., & Subrahmanian, V. S. (2007, March). Sentiment analysis: Adjectives and adverbs are better than adjectives alone. In ICWSM

Author's Profile

Rahim Dehkharghani: he received his PhD in Computer Science and Engineering, Artificial Intelligence branch, from Sabanci Universty, Istanbul, in 2015 and his MSc degree in Computer Engineering, Software branch from Shahid Beheshti University, Tehran, in 2007. His main research area is Natural Language Processing and Data Mining. He specifically works on sentiment analysis of textual data. He is a faculty member of Computer Engineering group at University of Bonab (Bonab, Iran) since 2016.