2019 IEEE International Conference on Big Data (Big Data)

A Rule-Based Holistic Approach for Turkish Aspect-Based Sentiment Analysis

Kivanc Bayraktar
Computer Engineering
Gazi University
Ankara, Turkey
bayraktarkivanc@gmail.com

Uraz Yavanoglu
Computer Engineering
Gazi University
Ankara, Turkey
uraz@gazi.edu.tr

Alper Ozbilen
Information Security Engineering
Gazi University
Ankara, Turkey
ozbilen@gmail.com

***Abstract*—In this study, a holistic method which uses statistical, linguistic and rule-based approaches for Turkish aspect-based sentiment analysis is proposed. The proposed method has been tested on the Turkish restaurant dataset created within the scope of SemEval Aspect Based Sentiment Analysis (ABSA) 2016. Firstly, candidate aspect terms were acquired employing LDA, C-value and WSBFE. Afterwards, aspect terms were found by rule-based approach and aspect-sentiment pairs were determined. In aspect term extraction 56,28% f-score was obtained while in aspect-sentiment matching phase 52,05% accuracy was achieved.**

Keywords—sentiment analysis, aspect-based sentiment analysis, Turkish sentiment analysis

I. Introduction

Opinions of people greatly impact the decision-making process of other people and institutions providing services to the society. Today, internet has become an integral part of our daily lives as people can easily disseminate their thoughts through the internet. Along with the rapidly increasing data sizes, data processing has become even more challenging. As a result, the concept of sentiment analysis has emerged. Sentiment analysis is divided into document level sentiment analysis, sentence level sentiment analysis and aspect-based sentiment analysis [1]. Document-level sentiment analysis aims to classify the overall sentiment in a document as positive, negative or neutral. Sentence level sentiment analysis aims to classify the overall sentiment in a sentence as positive, negative or neutral. Liu [1] stated that "both the document level and the sentence level analyses do not discover what exactly people liked and did not like."

Aspect-based sentiment analysis has two common steps: aspect term extraction and aspect term classification [2]. In aspect extraction phase, the subjects in the text are found. In aspect classification phase, aspects are classified as positive, negative or neutral. There are studies that use deep learning [3], [4], topic modeling [5], [6] and linguistic rule-based approaches [7] in aspect-based sentiment analysis.

In this study, a holistic approach has been proposed to achieve aspect term inference and aspect term classification in Turkish. To discover aspect terms LDA, C-value and WSBFE methods are utilized. To identify aspect-sentiment pairs rule-based approach has been suggested. Experimental study realized with SemEval [1] ABSA 2016 Turkish restaurant dataset.

Rest of the paper is organized as follows: the following section presents a literature survey related to the problem. Section 3 gives the brief idea about the utilized methods. Section 4 describes the proposed model. Section 5 contains experimental results and section 6 contains conclusions and future work.

II. Related Work

Turney [8] proposed a simple unsupervised learning algorithm to classify reviews as recommended or not recommended. The classification of the reviews was carried out using the mean semantic orientation value of phrases containing adjectives or adverbs. Various rules have been used to find phrases. It is stated that the phrase has positive semantic orientation in case of good association and negative semantic orientation in case of bad association. In the study, the semantic orientation of a phrase was calculated by subtracting the PMI value between the phrase and the word "excellent" from the PMI value between the given phrase and the word "poor". The document is classified as recommendable or not, depending on whether the sum of the semantic orientation values of the phrases is greater than zero.

A problem of creating a feature-based summary from comments on products sold online was addressed by Hu and Liu [9]. Common features were identified by using association rule mining (apriori). Adjectives closest to these features were accepted as opinion words. A list of words expressing emotion was created using these opinion words and a specific root list. The opposite and synonym words of these words were added to this list, and the semantic orientations of the words in the list were found with the help of WordNet. Features not found by association rule mining were obtained by finding nouns/noun groups close to the opinion words. Afterwards, each sentence was classified as positive or negative and a summary based on feature was produced. Correspondingly, Popescu and Etzioni [10] increased the precision of the target inference achieved by Hu and Liu [9] from 64% to 86%, by utilization of PMI method in determination of the aspect terms. Suggested aspect-based sentiment analysis method was named as *OPINE*. Relaxation labeling method was used to calculate semantic orientation.

Another system developed for domain-independent feature extraction, namely *WhatMatter*, was proposed by Siqueira and Barros [11]. This system consists of 4 stages: (i) frequent nouns identification, (ii) relevant nouns identification, (iii) feature indicators mapping, (iv) unrelated nouns removal. In the first stage, nouns in the text are found and those above the threshold are called as candidate features. In the second stage, adjectives next to the candidate features were found and nouns close to those adjectives are added to the candidate features. In the third stage, domain-dependent manually generated feature-opinion pairs were used to find the implied features. In the last stage, PMI technique was used to remove irrelevant features.

[1] http://alt.qcri.org/semeval2016/task5/

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE
978-1-7281-0858-2/19/$31.00 ©2019 IEEE

Another study, in which the dictionary created by Hu and Liu [9] was employed as the initial dictionary, was Qiu et al. [12]. In scope of the study, the problems of expanding the opinion dictionary and extracting the aspect terms were examined. Some syntactic relations between the opinion words and the aspects were found with the help of dependency parser, and these relationships were used to expand the initial opinion dictionary and to find aspect terms. With the double propagation approach, other new opinion words and aspects were found by using the initial opinion dictionary. This cycle continues until new aspect and opinion words cannot be found.

Most of the research in the sentiment analysis field focuses on English. Recently, there is an increase in the number of studies on sentiment analysis on Turkish. Coban et al. [13] tagged the Turkish data set obtained from Twitter as positive or negative by using pre-determined feeling symbols. After the preprocessing steps, the features were extracted using the BoW and n-gram models. TF, Boolean and TF-IDF methods were used for the weighting of the extracted features, CfsSubset algorithm was applied for the feature selection. NB, multinominal Naïve Bayes (MNB), SVM and k-NN were used for classification and the results were compared. MNB gave the most successful result in BoW and n-gram models.

Oğul and Ercan [14] created a Turkish hotel review dataset. They performed sentiment analysis using MNB, RF, SVM and dictionary based SentiTFIDF methods. RF is the most successful method with 82% sensitivity.

WSBFE (Web Search Based Feature Extraction), a domain-independent and unsupervised feature extraction method for feature-based sentiment analysis in Turkish texts, was proposed by Kama et al. in [2]. With this method, it is aimed to increase the performance of frequency-based feature extraction using search engine. First phase, nouns and noun groups in the text are found then which above certain value is passed to feature list. A search query is created for each element in the list to be sent to the search engine API. After these queries are sent to the Web search engine, the number of occurrences of these search results are stored in the list. Second phase, nouns and noun groups with their occurrence counts is sorted in descending order. Nouns and noun groups with frequencies above this threshold are returned as features. Third phase, new features are extracted from noun groups that are below the threshold. For each such noun group, it is checked whether at least one word of this noun group belongs to previously returned features list or not. If at least one such word exists, then this noun group is also added to the list.

Bilgin et al. [15] developed the Turkish WordNet by translating the synonyms sets in WordNet within the scope of the Balkanet project. Dehkharghani et al. [16] created first Turkish polarity source, name as SentiTurkNet. They assigned positivity, negativity and neutrality scores to each synonym set in Turkish WordNet. Ucan et al. [17] proposed an automatic translation-based approach to construct sentiment lexicon for different languages using SentiWordNet. They obtained 3 different dictionaries for Turkish.

III. BACKGROUND

A. Latent Dirichlet Allocation (LDA)

LDA is a generative probabilistic model of a corpus. The basic idea is that documents are represented as random mixtures over latent topics, where each topic is characterized by a distribution over words [18]. LDA is a topic modeling method utilizing bag of words (BoW) to uncover hidden topics within the dataset [18]. The words with the highest probability on each topic give a good idea of what the topic might mean [19].

BoW increases the efficiency, however, causes loss in inter-word semantic relationship information. To eliminate the BoW constraint, the dataset was updated using frequently used noun groups, then the LDA was then run.

B. Pointwise Mutual Information (PMI)

PMI is used to measure association between two terms [20]. This formula given below.

$$\mathrm{PMI}(\mathrm{word}_1,\mathrm{word}_2)=\log_2\frac{\mathrm{P}(\mathrm{word}_1,\ \mathrm{word}_2)}{\mathrm{P}(\mathrm{word}_1)\mathrm{P}(\mathrm{word}_2)} \quad (1)$$

$\mathrm{P}(\mathrm{word}_1, \mathrm{word}_2)$ represents the probability of coexistence of word_1 and word_2, while $\mathrm{P}(\mathrm{word}_1)\mathrm{P}(\mathrm{word}_2)$ refers to the probability that the two terms coexist when they are statistically independent.

Turney used the following equations to calculate the PMI score [8].

$$\mathrm{P}(\mathrm{word}) \equiv \mathrm{hits}(\mathrm{word}) \quad (2)$$

$$\mathrm{P}(\mathrm{word1},\ \mathrm{word2}) \equiv \mathrm{hits}(\mathrm{near}(\mathrm{word1},\ \mathrm{word2})) \quad (3)$$

hits(X) refers to number of times the given word X is given, near(Y, Z) refers to noun groups with a certain distance, beginning with Y and ending with Z. This hit function is used in the WSBFE approach.

C. C-value

C-value is a domain-independent method used for multi-word aspect term recognition [21]. This method aims to find candidate multi-word terms from the dataset using various filtering rules.

Ventura et al. [22] modified the C-value method to find one-word terms. It was aimed to eliminate the zero (0) coefficient problem for one-word terms by adding one (1) before calculating the logarithm value. C-value score can be calculated as follows:

- If a is not nested term

$$\text{C-value}(a)=\log_2|a+1|.f(a) \quad (4)$$

- otherwise

$$\text{C-value}(a)=\log_2|a+1|.\ (f(a)-\frac{1}{P(T_a)}\ \sum_{b\in T_a} f(b)) \quad (5)$$

where a refers candidate string, $|a|$ refers count of string, $f(a)$ refers frequency of occurrence in dataset, T_a refers set of extracted longer candidate terms that contain a, $P(T_a)$ refers number of distinct Ta, $\sum_{b\in T_a} f(b)$ refers total frequency by which a appears in longer strings.

IV. Proposed Methodology

Dataset is separated into sentences and the spelling errors are corrected with the help of the natural language processing library Zemberek[2] and Yandex.XML[3]. After that each word in the dataset is converted to lowercase and the inflectional suffixes of the words are removed. The main reason for removing inflectional suffixes instead of finding the root of each words is to prevent the words from losing their meaning.

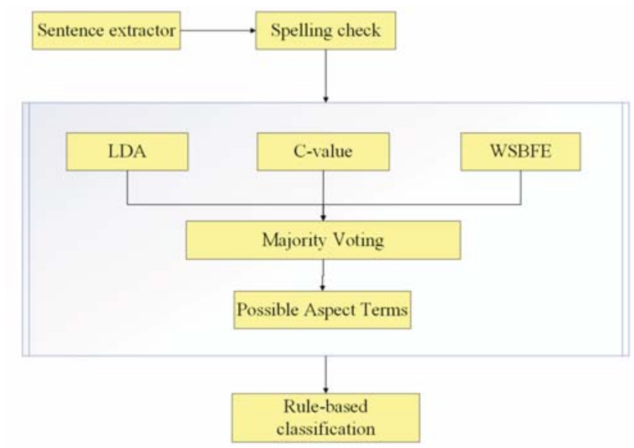

Fig. 1. Flowchart of the proposed method

There are approaches that use ensemble of classifiers in document level sentiment analysis [23], [24]. In aspect extraction phase, ensemble of some feature extraction methods utilized. LDA, WSBFE and C-value techniques are used to find the possible aspect terms. The expressions found by at least two methods are accepted as the aspect term (majority vote).

A rule-based approach is proposed to find real aspects and classify aspect-sentiment pairs using the target terms and two different emotion dictionaries. The rule-based approach takes possible terms and sentences as input and gives the real aspect term list and the aspect-sentiment pairs as output. The overall methodology is shown in Fig. 1.

Firstly, opinion dictionary was obtained by automatically translating the dictionary created by Hu and Li [9] to Turkish using Tureng[4] Dictionary. Verbs and nouns expressing opinions were added to this dictionary and the dictionary was expanded. Secondly, dictionary is SentiTurkNet [16], developed by Dehkharghani et al. emotion expressions are primarily searched in the first dictionary, and expressions not found in the first dictionary are searched on SentiTurkNet. Thus, it is aimed to determine the polarity of the term.

The set of rules implemented within the algorithm is presented in Fig. 2 as a holistic, closed logical expression. With the rule-based approach, it is expected that the statements that conform to this rule will be solved. In this approach, conjunctions like "*ama*", "*fakat*" are not taken into consideration. In this notation I: represents the names and names groups as the aspect term, Z: adverbs, S: the adjectives found in the opinion dictionary, F: the verbs in the opinion dictionary. "*" no expression or repeated occurrence, "?" expression never found or found once, "|" expression or means. In TABLE I. English conversions of Turkish expressions are given.

```
I (('ve'|'ile'|',') I) (Z)* (S |F ('de'|'da')? ('değil')?) (('ve' | ',') ((Z)* | F ('de'|'da')? ('değil')?))*
```

Fig. 2. Logical expression

TABLE I. Turkish Expressions

Turkish expression	**English conversion**
ama, fakat	but
ve	and
ile	with
de-da	also
değil	not

V. Experimental Results

Turkish restaurant dataset created within the scope of SemEval ABSA 2016 was used to evaluate performance of proposed methods. This dataset contains 1232 training and 144 test sentences. 128 of the training sentences and 20 of the test sentences are out of scope, meaning that there are no emotion definition groups. The sample part of the dataset is given in Fig. 3.

SemEval ABSA 2016 competition has two modes, constrained and unconstrained. While constraint system produces results using only the data set provided by the committee, unconstrained system produces results by using external data. Since our work is a closed system, the training set was used only during the aspect term discovery phase. F-score was used to evaluate aspect term extraction in scope of competition. Opinion groups with the target term "NULL" were excluded from the evaluation.

```
<Review rid="5002">
  <sentences>
    <sentence id="5002:0">
      <text>Menü çeşitleri zengin.</text>
      <Opinions>
        <Opinion target="Menü çeşitleri" category="FOOD#STYLE_OPTIONS" polarity=
        "positive" from="0" to="14"/>
      </Opinions>
    </sentence>
    <sentence id="5002:1">
      <text>Fiyatları normalin biraz üstünde ama yinede caddedeki yorgunluğu atmanız
      için ferah ferah oturabileceğiniz mekan.</text>
      <Opinions>
        <Opinion target="mekan" category="AMBIENCE#GENERAL" polarity="positive" from=
        "107" to="112"/>
        <Opinion target="Fiyatları" category="RESTAURANT#PRICES" polarity="negative"
        from="0" to="9"/>
      </Opinions>
    </sentence>
  </sentences>
</Review>
```

Fig. 3. Example of the dataset

In baseline work [25], distinct dictionaries were created for all target categories in the training data. All terms in the training set representing categories were filled into these dictionaries. Then aspect category was assigned to the sentences with the linear core SVM algorithm. Since the category was known, the terms in the sentence which belongs to category are identified as the aspect term. In sentiment polarity detection step a Support Vector Machine (SVM) with a linear kernel was used. Cetin and Eryigit [26] proposed sequence labeling algorithm based on conditional random fields (CRF) for aspect category identification and aspect term extraction. Then a linear classification method based on feature selection from positionally and syntactically neighboring tokens were applied. INSIGHT [27] team proposed a model based on the deep learning-based approach

[2] https://github.com/ahmetaa/zemberek-nlp
[3] https://tech.yandex.com.tr/xml/
[4] https://tureng.com/tr/turkce-ingilizce

to aspect-based sentiment analysis, which employs a convolutional neural network for aspect extraction and sentiment analysis.

Possible aspect terms of the restaurant domain were determined by using LDA, WSBFE and C-value (majority voting). 284 possible terms have been found, of which 158 are related to this field.

The aspect term extraction approach was compared with the study using the SemEval ABSA 2016 Turkish dataset and the basic system presented by the committee. The results are shown in TABLE II. . According to these results, it can be seen that the rule-based approach applied by giving a list of possible aspect terms is better than the other studies.

TABLE II. ASPECT EXTRACTION RESULT

Work	F-score
Baseline [25]	41,86
Cetin & Eryigit [26]	53,12
Our Work	56,28

In SemEval ABSA 2016, the number of aspect categories those estimated correctly were divided by the total number of categories to calculate the accuracy of the aspect sentiment mapping. "NULL" terms were excluded from evaluation. Aspect category detection was not performed in this work. For this reason, to evaluate this system, the number of aspect terms whose polarity was correctly estimated were divided by the total number of target terms. There are 159 aspect-sentiment pairs and in 13 of them aspect term is "NULL".

With the rule-based approach, there are 110 different aspect-sentiment pairs, of which 76 are correctly classified. A comparison table of sentiment classification is given in TABLE III. .

TABLE III. ASPECT SENTIMENT MAPPING RESULT

Work	Accuracy
Baseline [25]	72,33
Cetin & Eryigit [26]	76,10
Insight [27]	74,21
Our work	52,05

The main reason of the low-level classification is that unpredicted aspect-sentiment pairs are considered as a misclassified. Another reason is that the rule-based approach does not propose a solution for idioms. Also, lack of attention to grammar rules is another factor.

VI. CONCLUSIONS AND FUTURE WORK

A holistic method which uses statistical, linguistic and rule-based approaches for Turkish aspect-based sentiment analysis is proposed in this work. Firstly, possible target terms were obtained by majority voting by using LDA, C-value and WSBFE methods. With rule-based approach real aspects and aspect-sentiment pairs are found. This system has higher aspect extraction f-score value than other methods but aspect-sentiment mapping accuracy lower than other methods.

The success of the rule-based approach varies depending on the possible aspect terms given as input and the capacity of the natural language processing library. As a result of Turkish is agglutinative, in rare situations, our method does not work accurately while finding the lemmas of the words.

Our method uses a general-purpose dictionary. Instead of using a general-purpose dictionary, using domain-specific dictionary will give better result. Also, the double propagation method can be used to increase the success of the aspect-sentiment matching of our approach.

REFERENCES

[1] B. Liu, "Sentiment analysis and opinion mining," *Synthesis lectures on human language technologies*, vol. 5, no. 1, pp. 1-167, 2012.

[2] B. Kama, M. Ozturk, P. Karagoz, I. H. Toroslu, and O. Ozay, "A web search enhanced feature extraction method for aspect-based sentiment analysis for turkish informal texts," in *International Conference on Big Data Analytics and Knowledge Discovery*. Springer, 2016, pp. 225-238.

[3] Y. Ma, H. Peng, T. Khan, E. Cambria, and A. Hussain, "Sentic lstm: a hybrid network for targeted aspect-based sentiment analysis," *Cognitive Computation*, vol. 10, no. 4, pp. 639-650, 2018.

[4] M. Al-Smadi, B. Talafha, M. Al-Ayyoub, and Y. Jararweh, "Using long short-term memory deep neural networks for aspect-based sentiment analysis of arabic reviews," *International Journal of Machine Learning and Cybernetics*, vol. 10, no. 8, pp. 2163-2175, 2019.

[5] A. García-Pablos, M. Cuadros, and G. Rigau, "W2vlda: almost unsupervised system for aspect based sentiment analysis," *Expert Systems with Applications*, vol. 91, pp. 127-137, 2018.

[6] R. K. Amplayo, S. Lee, and M. Song, "Incorporating product description to sentiment topic models for improved aspect-based sentiment analysis," *Information Sciences*, vol. 454, pp. 200-215, 2018.

[7] A. S. Shafie, N. M. Sharef, M. A. A. Murad, and A. Azman, "Aspect extraction performance with pos tag pattern of dependency relation in aspect-based sentiment analysis," *in 2018 Fourth International Conference on Information Retrieval and Knowledge Management (CAMP)*. IEEE, 2018, pp. 1-6.

[8] P. 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*. Association for Computational Linguistics, 2002, pp. 417-424.

[9] M. Hu and B. Liu, "Mining and summarizing customer reviews," in Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2004, pp. 168-177.

[10] A.-M. Popescu and O. Etzioni, "Extracting product features and opinions from reviews," in *Natural language processing and text mining*. Springer, 2007, pp. 9-28.

[11] H. Siqueira and F. Barros, "A feature extraction process for sentiment analysis of opinions on services," in *Proceedings of International Workshop on Web and Text Intelligence*, 2010, pp. 404-413.

[12] G. Qiu, B. Liu, J. Bu, and C. Chen, "Opinion word expansion and target extraction through double propagation," *Computational linguistics*, vol. 37, no. 1, pp. 9-27, 2011.

[13] O. Coban, B. Ozyer, and G. T. Ozyer, "Sentiment analysis for turkish twitter feeds," in *2015 23nd Signal Processing and Communications Applications Conference (SIU)*. IEEE, 2015, pp. 2388-2391.

[14] B. B. Oğul and G. Ercan, "Sentiment classification on turkish hotel reviews," in *2016 24th Signal Processing and Communication Application Conference (SIU)*. IEEE, 2016, pp. 497-500.

[15] O. Bilgin, O. Cetinoğlu, and K. Oflazer, "Building a wordnet for turkish," *Romanian Journal of Information Science and Technology*, vol. 7, no. 1-2, pp. 163-172, 2004.

[16] R. Dehkharghani, Y. Saygin, B. Yanikoglu, and K. Oflazer, "Sentiturknet: a turkish polarity lexicon for sentiment analysis," *Language Resources and Evaluation*, p. 1-19, 2016.

[17] A. Ucan, B. Naderalvojoud, E. A. Sezer, and H. Sever, "Sentiwordnet for new language: Automatic translation approach," in *2016 12th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*. IEEE, 2016, pp. 308-315.

[18] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *Journal of machine Learning research*, vol. 3, no. Jan, pp. 993-1022, 2003.

[19] H. Jelodar, Y. Wang, C. Yuan, X. Feng, X. Jiang, Y. Li, and L. Zhao, "Latent dirichlet allocation (lda) and topic modeling: models, applications, a survey," *Multimedia Tools and Applications*, pp. 1-43, 2018.

[20] K. W. Church and P. Hanks, "Word association norms, mutual information, and lexicography," *Computational linguistics*, vol. 16, no. 1, pp. 22-29, 1990.

[21] K. Frantzi, S. Ananiadou, and H. Mima, "Automatic recognition of multiword terms:. the c-value/nc-value method," *International journal on digital libraries*, vol. 3, no. 2, pp. 115-130, 2000.

[22] J. A. Lossio-Ventura, C. Jonquet, M. Roche, and M. Teisseire, "Combining c-value and keyword extraction methods for biomedical terms extraction," in *LBM: Languages in Biology and Medicine*, 2013.

[23] N. F. Da Silva, E. R. Hruschka, and E. R. Hruschka Jr, "Tweet sentiment analysis with classifier ensembles," *Decision Support Systems*, vol. 66, pp. 170-179, 2014.

[24] C. Catal and M. Nangir, "A sentiment classification model based on multiple classifiers," *Applied Soft Computing*, vol. 50, pp. 135-141, 2017.

[25] M. Pontiki, D. Galanis, H. Papageorgiou, I. Androutsopoulos, S. Manandhar, A.-S. Mohammad, M. Al-Ayyoub, Y. Zhao, B. Qin, O. De Clercq et al., "Semeval-2016 task 5: Aspect based sentiment analysis," in *Proceedings of the 10th international workshop on semantic evaluation (SemEval-2016)*, 2016, pp. 19-30.

[26] F. S. Cetin and G. Eryigit, "Turkce hedef tabanlı duygu analizi icin alt gorevlerin incelenmesi-hedef terim, hedef kategori ve duygu sınıfı belirleme," *Bilisim Teknolojileri Dergisi*, vol. 11, no. 1, pp. 43-56, 2018.

[27] S. Ruder, P. Ghaffari, and J. G. Breslin, "Insight-1 at semeval-2016 task 5: Deep learning for multilingual aspect-based sentiment analysis," in *Proceedings of the 10th international workshop on semantic evaluation (SemEval-2016)*, 2016, pp. 330-336.