Applied Soft Computing 118 (2022) 108524

Contents lists available at ScienceDirect

Applied Soft Computing

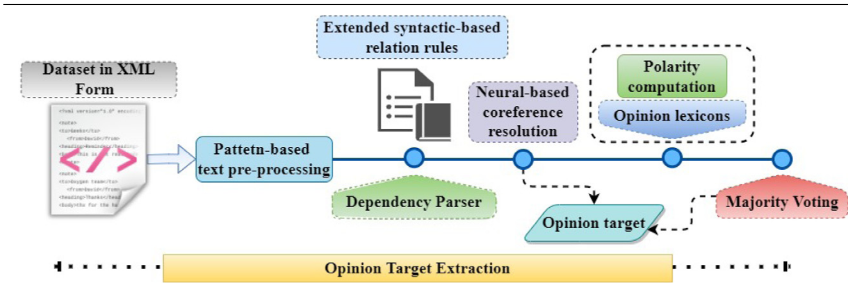
journal homepage: www.elsevier.com/locate/asoc

Extended rule-based opinion target extraction with a novel text pre-processing method and ensemble learning

Kürşat Mustafa Karaoğlan *, Oğuz Fındık

Department of Computer Engineering, Karabuk University, Karabük 78000, Turkey

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 5 October 2021
Received in revised form 3 January 2022
Accepted 23 January 2022
Available online 1 February 2022

Keywords:
Opinion mining
Aspect-based sentiment analysis
Opinion target extraction
Aspect extraction
Pattern-based text pre-processing
Syntactic-based relation rules
Ensemble learning
Unsupervised learning

ABSTRACT

Opinion target extraction (OTE) is the extraction of explicit expressions related to entity aspects interpreted with subjective attributive words in the review sentences using supervised or rule-based approaches. Despite the constraints of syntactic-based relation rules, rule-based approaches can be domain-independently implemented. Although supervised approaches yield better results, more costly due to requiring a large number of labeled samples. This study proposes an unsupervised (rule-based) OTE approach with novel methods and extended rule-based techniques to overcome the aforementioned issues. In this study, first, a novel pattern-based text pre-processing method is proposed to eliminate punctuations that are incompatible with determinative group rules patterns. Then, implemented syntactic-based relation rules on the dependency relation graph are extended with new auxiliary features to extract multi-word expressions which modify each other. The majority voting method is used for optimizing the performance of outputs. Finally, the effectiveness of the proposed approach was tested on a restaurant review dataset. The experimental results show that the proposed approach outperforms all unsupervised approaches. Additionally, it gives comparable results with the supervised approaches, revealing the effectiveness of the proposed approach.

1. Introduction

Currently, web platforms provide an opportunity for businesses that make online marketing to expand their trading [1]. As a result of this, companies face large user populations [2]. Consumers can express their subjective opinions about any product, service, or employee [3]. This situation has constituted vast volumes of unstructured text sources with feedback content such as consumer-generated blogs and product reviews [4]. On the other hand, organizations and researchers focus on opinion analysis of these unstructured text data sources to determine how users evaluate products to improve the service quality of businesses [5, 6]. In these analyses, manual data processing is not practical to perform opinion analysis due to the vast dataset size [7]. Hence, opinion analyses are performed by automated approaches based on a combination of Data Mining, Web Mining, and Text Mining techniques [6].

In Text Mining, Sentiment Analysis (SA), or Opinion Mining (OM) is used chiefly for similar purposes about opinion extractions [8]. Researchers generally deal with OM at document,

* Corresponding author.
E-mail addresses: kkaraoglan@karabuk.edu.tr (K.M. Karaoğlan), oguzfindik@karabuk.edu.tr (O. Fındık).

https://doi.org/10.1016/j.asoc.2022.108524
1568-4946/

sentence, and aspect levels [9]. Document and sentence-level strategies deal with the general polarity of texts. For this reason, it is not possible to measure how users evaluate the subject of the text [10]. A more complex study is required in aspect-level strategies than associating the document to general polarity. Besides, aspect-level strategies offer a more realistic approach, considering the potential of polarities as independently within review texts containing subjective opinions [11]. Therefore, fine-grained analysis about user opinions is performed at the aspect level. These analyses examine the relations between opinion target expressions (opinion targets) and subjective attributive (opinion words) computed polarity values. Opinion target is an explicit expression of aspect within the text of a subjective opinion that describes an entity [12]. Opinion target extraction (OTE) extracts this explicit expression with which the subjective opinions are related, with the help of the polarity values of the subjective (positive or negative) opinion words in the texts [10].

OTE approaches generally use supervised or rule-based models. Although successful results are obtained using supervised models, they need a lot of labeled data for training [13]. On the other hand, unsupervised rule-based approaches can be integrated into systems regardless of the domain with syntactic-based relation rules [14].

This study is aimed to automatically extract the opinion target expressions of the aspects of the entity interpreted with subjective attributive words from the review texts. Thus, businesses are provided with an opportunity to analyze the review sentences and identify the product features that affect consumers' purchasing preferences. For this purpose, in this study, the unsupervised (rule-based) OTE approach is proposed to extract explicit expressions associated with subjective opinion words on review texts. In the proposed approach, first, a pattern-based pre-processing (PBP) method was developed on unstructured input texts to improve OTE performance, instead of traditional Natural Language Processing (NLP) pre-processing techniques (such as removing stopwords, stemming, lemmatization). Thanks to the PBP method, only punctuation marks that are not suitable for determinative group grammar rules in the input texts are removed with pattern-based rules, and input texts are corrected while preserving their meaning and structure. Thus, we intact other punctuation marks and all stopwords. Then, opinion target extraction was performed related to subjective opinion words on the parsed relation dependency graph with syntactic-based rules. Implemented grammar rules have been extended by adding new features so that it can capture multi-word expressions. In addition, the neural-based coreference resolution (NBCR) model, which has a higher performance than other models, was used to detect words with the same reference in input texts. Thus, the NBCR model outputs and opinion targets obtained were compared, and correct expressions that could be written interchangeably were determined. Finally, the majority voting (MV) method is implemented to optimize the output performances in the proposed OTE approach.

The experiments and performance evaluation of the proposed OTE approach were accepted by the researchers as a benchmark dataset (OTE dataset [13]) and were carried out using International Workshop on Semantic Evaluation (SemEval)- Aspect-Based Sentiment Analysis (ABSA) restaurant dataset. According to experimental results obtained, the effectiveness of the proposed OTE approach was observed. While better performance was obtained from the proposed method than other rule-based methods, comparable results were obtained with supervised methods.

The main contributions of the study are as follows:

- A novel PBP method is presented to improve OTE performance for ABSA tasks. With this method, instead of general NLP pre-processing techniques requiring handcraft engineering, only punctuation marks that are not suitable with determinative group grammar rules are removed from the texts using Dependency Parsing. In addition, another contribution of this method is intact of the other original punctuation marks and stopwords that contribute to the meaning, structure, and analysis of the sentences.
- Based on syntactic-based rules of OTE, new auxiliary relation types components are integrated into the fundamental rule structure that analyzes parsed dependency relation graph to extract multi-word expressions which modified each other by the rule-based structure.
- Performance optimization was carried out by combining the outputs of OTE models with the MV method, which is one of the ensemble learning strategies.

The rest of the paper is structured as follows: Section 2 presents the existing related works for ABSA and OTE tasks. Section 3 presents the developed PBP, extended rule-based algorithms for opinion target extraction, and MV methodologies after an overview of the proposed approach. Section 4 covers the OTE dataset we use for evaluation, opinion lexicons, performance metrics, and performed experimental results with performance comparison. We present the discussion in Section 5 and the conclusion and future works in Section 6.

2. Related work

OM is studied under the study areas of information retrieval, natural language processing, and text mining [15]. In OM, aspect-based (fine-grained [16]) text analysis on documents, sentences, and entities are performed [17]. For this purpose, OM generally focuses on problems such as polarity determination and classification, subjectivity classification, opinion summarization, document/sentence-level sentiment analysis, and ABSA [18]. Document [19] and sentence-based [20] analysis provide functionally useful results within texts containing opinions [8]. However, considering the independent opinions within text content, it is necessary to work at the aspect-level, which provides fine-grained analysis to obtain details [6].

ABSA aims to identify subjective information within texts, extract them, classify the extracted information intellectually based on polarity, and analyze them at various levels [15]. ABSA is studied by researchers with titles such as sentiment analysis [21], opinion extraction [12], sentiment mining, subjectivity analysis [22], affect analysis, emotion analysis [23], review mining, or review analysis [24]. Studies about ABSA focus on major tasks like OTE [13,25], Aspect Category Detection [26] and lexicon-based opinion analysis [8] / polarity detection. OTE strategies provide more effective and fine-grained analysis to extract entities with independent poles in texts [5,27].

In general, OTE studies have included Supervised, Unsupervised, and Semi-Supervised methods [8] that compose of Rule-Based approaches [10,28], Machine Learning [29], Deep Learning [12], or Hybrid models [5,13]. First studies on OTE presented tokens having Part of Speech (POS) features within texts or approaches using the frequency-dependent properties of extractions supported by the dependency relation. These frequency-dependent extractions caused constraints to detect infrequent valuable terms [30]. For OTE, Conditional Random Field (CRF) [31], Hidden Markov Model (HMM), Support Vector Machine (SVM) based approaches were proposed in the first supervised approaches of machine learning using data whose inputs are created by sequential learning. In CRF approaches, OTE performance has been enhanced by collocations of Long Short-Term Memory (LSTM), Bidirectional-LSTM [19], and Recurrent Neural Network (RNN) [32] models. Semi-supervised approaches are given in the literature that uses statistical models, Maximum-Entropy and Latent Dirichlet Allocation (LDA), including seed

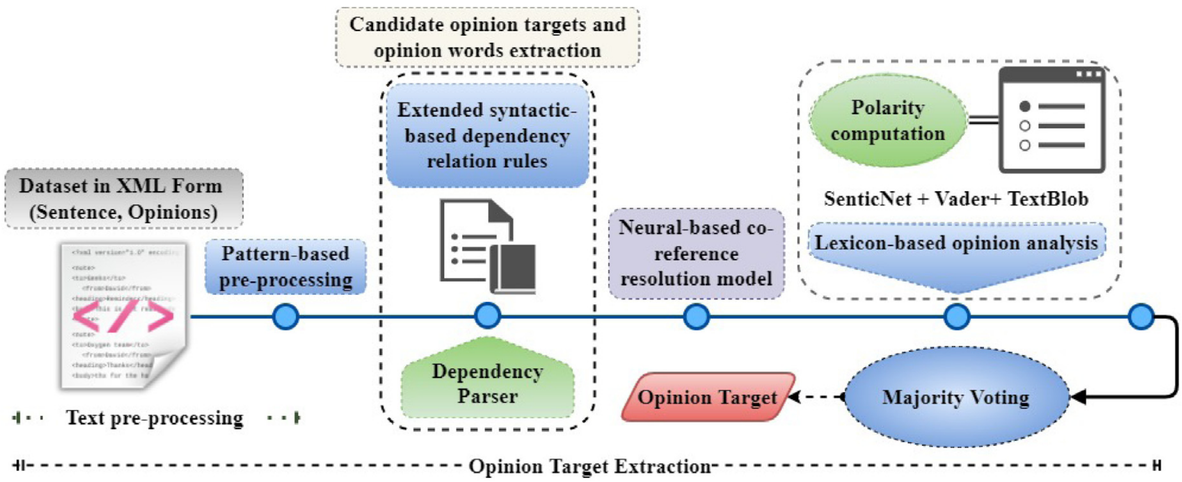


Fig. 1. An overview of the proposed OTE approach.

words for OTE [33]. Deep Learning approaches like Convolutional Neural networks (CNN) and RNN have successfully been applied for OTE tasks [34]. In recent years, evolutionary approaches for OTE tasks and approaches that automatically learn rules and generate general OTE linguistic patterns have also been included in the literature [35].

In unsupervised approaches, they assume that names within texts as potential aspects [14], the frequency values of each potential aspect have been calculated. These frequency values have been pruned according to specific threshold values. The adjective words closest to the extracted opinion target have been accepted as opinion words. In further studies, instead of pruning after calculating frequency values, the use of Pointwise Mutual Information (PMI) [36] scores has increased the extraction accuracy. Rule-based approaches play an essential role for OTE, either standalone or successfully integrating with other approaches [28]. In rule-based approaches, dependency parsing-based methods depending on the grammatical rules of the language are included in determining the relations between opinion targets and opinion words [14]. On the other hand, opinion word analysis in a lexicon-based (OALB) and OTE are complements of each other. Compared to the associated opinion words that describe different aspects, polarity computation and classification of opinion such as positive, negative, or neutral are tasks of OALB [21]. Dictionary-based [10] and corpus-based approaches [37] are presented for OALB tasks in the literature.

3. The proposed methodology

In this section, we first provide an overview of the OTE approach proposed in Section 3.1. Then, we describe the details of the developed PBP method structure in Section 3.2. In Section 3.3, we explain extended rule-based opinion target extraction algorithms. Finally, the implemented MV method will be presented in Section 3.4.

3.1. Overview

The architecture of the proposed OTE approach is shown in Fig. 1. This section explains the proposed approach by dividing it into three main topics: PBP, extended syntactic-based relation rules, and the MV method. In addition, the OTE dataset, OALB, and performance metrics shown in Fig. 1 are presented in detail under the experimental study section.

3.2. Text pre-processing method (PBP)

There may be many incompatible punctuation marks in the texts created by users on online platforms, and this situation may lead to limitations in the batch handling of the texts or advanced analysis. Especially in tasks such as part-of-speech (POS) tagging and Dependency Parsing, since the syntactic structure of the sentences is taken into account, the presence of misspellings and punctuation marks causes incomplete inferences. To overcome the problems mentioned above, instead of NLP's traditional text pre-processing [38], specific text pre-processing methods are needed, applied limitedly without disturbing the sentence's syntactic structure and semantic integrity. The PBP method (Fig. 2) ensures that only incompatible punctuation marks are eliminated between groups of words (such as noun phrases) according to the pattern of determinative group rules. This way aims to improve OTE performance by arranging the input data and removing text noise.

As seen in Fig. 2, with the PBP method, the stopwords that contribute to the structure and meaning of the sentence remain intact, while other punctuation marks are removed. For example, "Service- friendly and attentive". in the review sentence, the word "service" is an opinion target expression, and this expression is only possible by removing the punctuation mark in the phrase "service-friendly". A similar situation is "Ambiance-relaxed and stylish". which can be exemplified by the sentence. The word "ambiance" which is the opinion target in this sentence, is possible by removing the punctuation mark.

The algorithm of the PBP method developed in Algorithm 1 is presented. The "checkPunct" and "checkPattern" functions are given in Algorithm 1 to check the effect of existing punctuation marks on the sentence structure. Dependency parser outputs are checked according to the pattern structure determined by the "checkPattern" function.

Algorithm 1 Algorithm of the PBP method developed for text pre-processing

Initialize $S \in \mathcal{S}entences$
for each s_i in $\mathcal{S}$ **do**
  if checkPunct(s_i) and checkPattern(s_i) **then**
    removePunct(s_{ik})   ▹ k is the index of the punctuation mark
  end if
end for

3.3. Extended rules for opinion target extraction

This section describes opinion targets and opinion words extraction by the extended rule-based algorithms. As presented in

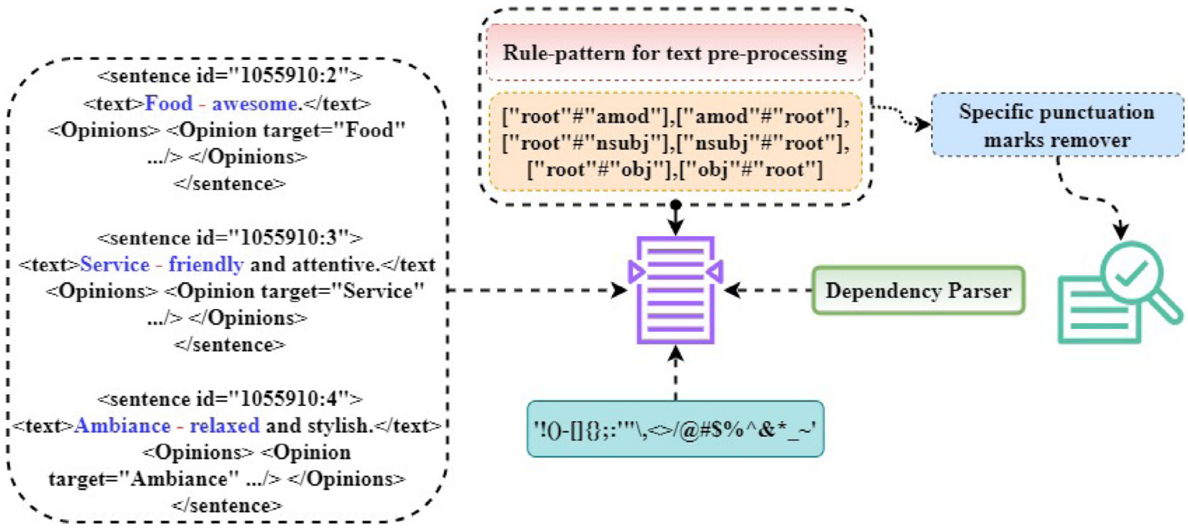


Fig. 2. PBP method with determinative group rule patterns.

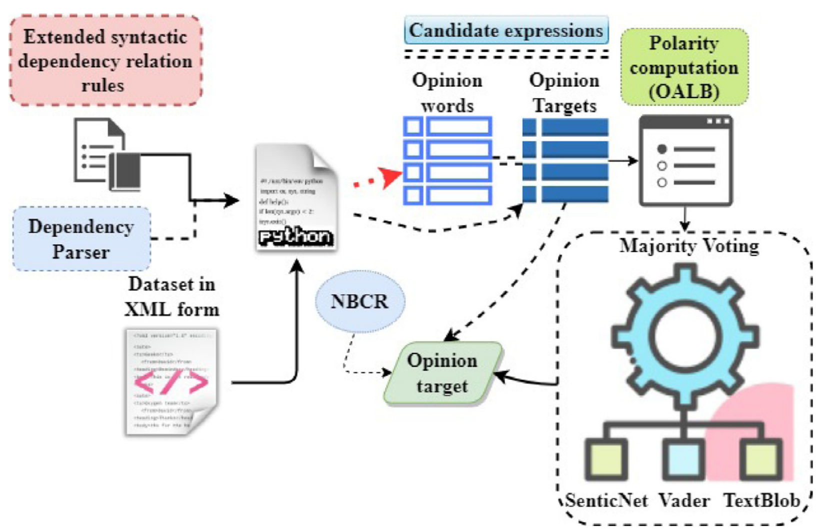


Fig. 3. The architecture of the proposed OTE approach with its components.

the diagram summarizing the architecture of the proposed OTE approach and components in Fig. 3, extended syntactic-based relation rules and data are used for inputs of this method.

The concept of syntactic-based relation rules is managed the asymmetrical bilateral relations between the lexical items of the syntactic structure. It connects the head (governor, superior, regent) and the dependent (modifier, inferior, subordinate) with the arrows presented in Fig. 4 to represent the individual dependency. These dependency relations generated for each sentence are explained as triples Dependency relation type (deprel), Governor (g), Dependent (d) on a graph.

The extraction of candidate opinion targets and their opinion words are determined by evaluating the triple explaining the dependency relations according to rule patterns. Candidate opinion targets and opinion words extraction are the first and most essential tasks for OTE [10]. Syntactic-based rules carry out OTE based on Open Knowledge Extraction strategies [25,39,40]. These rules analyze parsed dependency relation hierarchical graphs and extract candidate opinion targets and opinion words. However, suppose opinion targets and opinion words expressions consist of multi-word expressions which modified each other, as seen in Fig. 5. In that case, extending the rule structure with new features is needed to increase parsing performance.

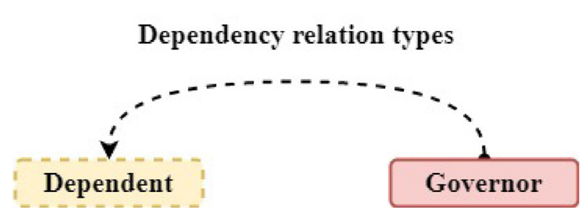


Fig. 4. The individual dependency representation between a governor and a dependency.

Dependency Parsing (Fig. 5) implemented in OTE is to construct a tree (graph) representing the relations between different words in a sentence. The relations between opinion words and opinion targets can be analyzed on the graph created after the dependency parser [39]. In OTE extractions, these expressions are considered potential opinion target candidates at every phase of the models.

Syntactic-based dependency relation structure developed with common relation types for the OTE task has been extended with auxiliary relation types. The relation types, explanations, and algorithms related to extended rules given in Table 1 are presented.

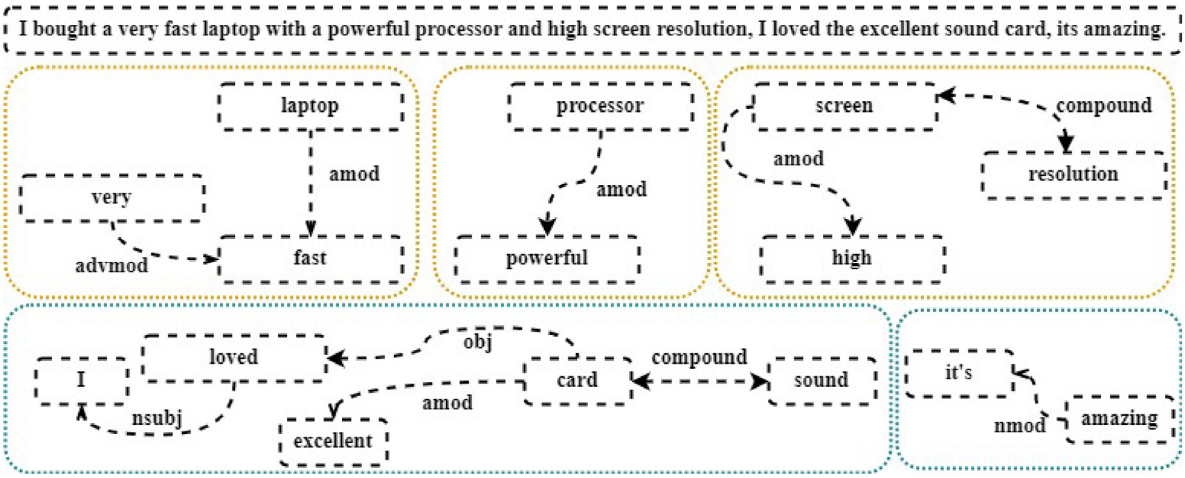


Fig. 5. The example of Dependency Parsing on a sentence including multi-word opinion target and opinion word.

Table 1
Syntactic-based dependency relation types [39] (d_i) and descriptions.

id	Relation types	Descriptions
d_1	**amod**	An adjective phrase used to change the meaning of noun phrases (adjectival modifier)
d_2	**nsubj**	Syntactic subject and the proto-agent of a clause (nominal subject)
d_3	**obj**	The noun phrase, which is the direct object of the verb, and its second-most core argument (object of verbs)
d_4	Compound	Compounds
d_5	advmod	An adverbial or adverbial phrase of a word (noun-clausal) that serves to modify a predicate or a modifier (adverbial modifier)
d_6	obl	Functionally corresponds to an adjective or an adverbial attached to a verb or another adverb (oblique nominal)
d_7	nmod	Noun functioning as oblique (non-core) or adjunct. Used for nominal modifiers of nouns or clausal predicates (nominal modifier)

In this context, auxiliary relation types such as compound, advmod, obl, and nmod have been added to fundamental dependency relation rules.

While $d_{1,2,3}$ in Table 1 are a fundamental relation types, others are auxiliary relation types to support the multi-word expressions process. According to the rules given in Algorithm 2, "ote" and "ow" components are selected from "g" and "dep" pairs with the help of defined fundamental and auxiliary relation types, using dependency parser function (deprel()) outputs.

Algorithm 2 The developed algorithm based on extended syntactic-based relation rules

Initialize $S^p \subset S \quad \wedge \quad e_i \in S_i^p$ ▷ S^p: opinion sentences
for each s_i in S_i^p **do**
  for each $deprel_j$ in deprel(s_i) **do**
    if $deprel_{ij}$=d_1 and compute(dep_{ij}) **then**
      $ote_{ij} \leftarrow multiWordOTE(g_{ij})$ ▷ ote: opinion target expression
      $ow_{ij} \leftarrow multiWordOW(dep_{ij})$ ▷ ow: opinion word expressions
    else if $deprel_i$=d_2 and compute(g_{ij}) **then**
      $ote_{ij} \leftarrow multiWordOTE(dep_{ij})$
      $ow_{ij} \leftarrow multiWordOW(g_{ij})$
    else if $deprel_i$=d_3 and compute(g_{ij}) **then**
      $ote_{ij} \leftarrow multiWordOTE(dep_{ij})$
      $ow_{ij} \leftarrow multiWordOW(g_{ij})$
    end if
  end for
end for

In the Algorithm 2, "e_i" represents the interpreted entity in the text, "ote_{ij}" represents the opinion targets of the entity and "ow_{ij}" represents the attributive words of opinion within a text. In addition, functions named deprel, compute, and multiWord are used. The deprel function is used to extract of all dependency relations for the sentence given as a parameter. Thus, desired dependency associations can be reached. Then, the Governor (g) and the Dependent (dep) pair, which create the dependency relation, are extracted for "ote" and "ow" assignments. (1) If the dependency relation type is d_1, the polarity value of the dep expression is calculated by the compute function. If it has the subjective value range, the "g" expression is continued as "ote". (2) For dependency relation types d_2 and d_3, the compute function calculate the polarity value of "g" expression. If it has the subjective value range, the "dep" expression is continued as "ote".

Finally, "g" and "dep" are discussed with modifier words. For expressions with two or more words, "ote" and "ow" are extracted with the developed multiWord function. $d_{1,2,3}$ relation types have a key role for extraction of multi-word terms in "multiwordrec()" function. In functions that whose outputs are multi-word expressions, "advmod" with $d_{1,2,3}$ and "obl" with d_1 relation types are used to extract multi-word opinion words. "nmod" relation type is used for the extraction of multi-word opinion targets with $d_{1,2}$.

OALB managed by the compute function, means the process of computation polarity on the opinion words associated with the candidate opinion target using linguistic resources such as SenticNet, Vader, and TextBlob (polarity analyzer models). Then, these linguistic resource outputs are used in the input of the MV.

After the opinion target expressions are extracted with rule-based algorithms in the OTE task, the same referenced expressions in the texts are extracted with the NBCR model. In this way, the results of the candidate opinion target are combined with the expressions that can be written instead of the obtained opinion target expression. The NBCR model can produce inefficient results in finding expressions with the same reference due to the complex structure of sentences. For this reason, parsing the NBCR model input sentences without disturbing the semantic integrity of complex sentences increases performance. For this reason, complex sentences in the dataset were parsed and converted into shorter sentences with sentence parsing. Fig. 6 presents actual, unparsed, and parsed the NBCR model outputs. As seen in the review sentence, the words referring to "the staff" can be determined correctly by the sentence parsing process. In this way, 'the staff' can be used instead of pronouns like "them" and "they" in the outputs.

Fig. 6. Tagged actual opinion target (in green at the top of the sentence) and same expression extraction by the NBCR model with (after) / without (before) sentence parsing process.

Table 2
Tabulated values of majority voting accuracy of L number of independent classifiers for individual accuracies [43].

Individual accuracies ($p > 0.5$)	Majority voting accuracy of L independent classifiers ($L = 3$)
0.6	0.6480
0.7	0.7840
0.8	0.8960
0.9	0.9720

3.4. Majority Voting (MV) method

Ensemble learning presents methodologies to use multiple classifiers, combine their classification predictions and select the classifier with the highest performance. Ensemble learning has a set of strategies used to combine the results of different classifiers and determine the most appropriate classifier for a particular problem [41]. These strategies combine results with majority voting (mean, a minimum, maximum, median, majority) [42,43] methods. The upper and lower accuracy limits of majority voting based on binary dependency can be derived by using the unique accuracy (p) results between the MV method with the prediction pool of independent classifiers (also called team, ensemble) and the classifiers with the number of classifiers as L. Let $D = \{D_1, \ldots, D_L\}$ be a set defined by classifiers $D_i : R^n \rightarrow \Omega$. Provided that $x \in R^n$, it is assumed that the class labels (x) defined in the classifier label set $\Omega = \{\omega_1, \ldots, \omega_n\}$ are assigned. Provided that the binomial formula $\Omega = \{\omega_1, \omega_2\}$, the general majority accuracy (P_{maj}) for individual p values are calculated among classifiers with an odd number of classifiers (L) by the binomial formula given in Eq. (1).

$$P_{maj} = \sum_{m=0}^{L/2} \binom{L}{m} p^{L-m}(1-p)^m \tag{1}$$

4. Experimental study

4.1. Experiment settings

4.1.1. Dataset

This study uses the English restaurant domain dataset in SemEval16-ABSA (OTE dataset) to test the proposed OTE approach. Most researchers accept this dataset as a benchmark dataset [5,13]. The significant factor in using this dataset, which is still state-of-the-art in the OTE approach, is that it has the right and already opinion target tags to evaluate the performance values obtained from the proposed approach. The dataset (Opinion Target, E#A, polarity) containing the restaurant reviews with more than one aspect consists of opinion target expressions, entity (E), and attributes (A) as implicit category, polarity, and other features. The other feature data consists of the sentence is, starting and ending offsets of opinion target expressions. Descriptive fundamental detailed information about the dataset and XML structure is presented in Table 3.

4.1.2. Opinion lexicons

In most OM applications, opinion lexicons are important in associating words with a polarity value. Commonly in ABSA, different levels of analysis can be performed by using these words, which have polarity values and are called opinion terms. Lexicon based methods utilize opinion lexicons to identify the sentiment orientation (polarity) scores or categories of these words [44]. Polarity values are obtained in a definite (positive, negative, neutral) or normalized numerical interval depending on the width and capability of the resource. SenticNet, a publicly available knowledge resource that uses artificial intelligence, linguistics, and psychology, includes more than 200,000 natural language concepts, related semantics, and sentence information, is preferred for polarity lexicon for concept-level sentiment analysis [45]. Although SentiWordNet inspired SenticNet, higher results were achieved based on performance in SenticNet. That is why SenticNet is preferred in this study. SentiWordnet generates outputs based on three numerical polarity scores (positive, negative, neutral) for each WordNet synset. Instead of category-dependent outputs, SenticNet generates polarity values in the interval of $[-1.0. +1.0]$. With the generated polarity values [10], it is possible to perform more flexible and complex analyses for OTE tasks.

As an alternative to SenticNet, TextBlob is an open-source library that can perform NLP methods and compute polarity values of sentences or words. A polarity value computed by TextBlob is derived by analyzer architecture based on Naïve Bayes [46]. TextBlob generates a pair of values in the form of "polarity, subjectivity" as output. Polarity and subjectivity scores are generated in the interval of $[-1.0. 1.0]$ and $[0.0. 1.0]$ respectively. While 0.0 for subjectivity score stands for "very objective", 1.0 stands for "very subjective". VADER (for Valence Aware Dictionary for sEntiment Reasoning) [47] is an MIT Licensed open-source software library that can work directly on unlabeled text data and take into account the density of polarities in addition to their classes (negative, neutral, positive). Thus, VADER provides a more sensitive analysis. It has been stated in the literature that VADER [48] shows an effective performance when it is used together with machine learning techniques like Naive Bayes,

Table 3
Descriptive statistics and XML structure for the OTE dataset.

Criteria	Count	XML structure of the OTE dataset
Reviews	350	
Sentences	2000	
Sentences with tagged opinion target	1708	
Total words	3086	<text> *&sentenceID - &sentence* </text> <Opinions> <Opinion target= *&OTE* category=***&E#A*** polarity=***&polarity*** from=***&startingoffset*** to=***&endingoffsets*** /> </Opinions>
Length of vocabulary	1999	
Parsable sentences	1771	
Total opinion targets	2507	
Range of word length in opinion targets (min. - max.)	1–15	
Single-word	1409	
Two-word	316	
Multi-word	507	
More than two-words	191	
More than three-words	72	

Maximum Entropy, and SVM and lexicons such as LIWC, ANEW, General Inquirer, SentiWordNet. In addition to polarity classes, it is possible to obtain a single class and score with the compound feature in VADER. Compound score [48] is computed by summing the rule-adjusted valance scores of each word in the dictionary and can be normalized in the interval of $[-1, +1]$. In this interval, -1 stands for "the most negative" and $+1$ stands for "the extreme positive". Researchers have suggested standard threshold values that divide the compound score into three categories as positive in the interval of [0.05,1.0], negative in the interval of $[-0.05, -1.0]$ and neutral in the interval of $(-0.05,+0.05)$.

4.1.3. Metrics

Our experiments use four metrics to evaluate the performance, i.e., precision (p), recall (r), F_1-score (f), and accuracy. Opinion targets are regarded to be correct if and only if they match the actual tags expressions. The micro-p, micro-r, micro-f, and accuracy are computed as follows:

$$micro - p = \frac{TP}{TP + FP} \tag{2}$$

$$micro - r = \frac{TP}{TP + FN} \tag{3}$$

$$micro - f = \frac{\mathrm{p} \times \mathrm{r}}{\mathrm{p} + \mathrm{r}} \tag{4}$$

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

4.2. Experimental results and performance comparison

OTE performance results of the proposed approach are obtained depending on the extended syntactic-based dependency relation rules algorithms by implemented PBP method at the input, NBCR model with sentence parsing, and MV methods. The performance results are obtained from the experimental studies performed on the OTE dataset that detailed, which is described in Section 4.1.1. The approach's success is evaluated by comparing the tagged opinion target expressions on the dataset.

The comparison results with other OTE baseline works are presented in Tables 4 and 5. According to performed analysis, the F-Measure score of the proposed OTE approach is measured as 0.70003. In Table 4, the results published in the proceeding book for SemEval16-ABSA Task 5 were reported based on F-Measure evaluation.

As can be seen from Table 4, the lowest F-Measure value is 0.50253. The average F-Measure value is 0.6271. The proposed approach is higher than the average F-Measure value. Besides, higher performance was obtained than the other 12 approaches reported based on the F-Measure value. The difference from AUEB approach is 0.00438. The performance of the proposed approach is very close to the most successful approach NLANGP on a comparable scale. Our model analyzes the performance of OTE with state-of-art models as well as recent approaches [5,10,13]. Table 5 presents relevant rule-based models and their F-Measure values.

As shown in Table 5, the lowest F-Measure value is 0.64240, and the highest F-Measure value is 0.66870. The average F-Measure value is 0.6522. The proposed approach obtained a higher F-Measure value of 0.70003 than other models. In the proposed OTE approach, the outputs of the three polarity models are combined by the MV method. Flowchart for calculations of both the combinational accuracy scores of the models and accuracy scores related to each polarity model are presented in Fig. 7. Besides, the obtained overall accuracy score after applying the MV method is presented on the graph in Fig. 7.

As shown in Fig. 7 the accuracy score of the proposed OTE approach after the MV method is 0.94342. With this method, a higher accuracy value has been achieved. It was concluded that the accuracy score obtained from the proposed approach is consistent with the values in Table 2 calculated using the formula ($p = 0.9$) given in Eq. (1).

5. Discussion

Our study showed that PBP, extended dependency relation rules, and MV methods contributed to the OTE performance according to the experimental results. These methods implemented in the proposed OTE approach have not been previously assessed. In this study, three new significant results were obtained. The first finding is that implementing the PBP method, integrated into all review analyses as domain-independent to the raw review texts, contributes to the OTE extraction performance. With this developed method, only punctuation marks that do not suitable with the determinative group grammar rules are removed without using an extra-linguistic resource, with dependency parsing. Thus, the original punctuation marks and stop words are preserved, keeping the meaning and structure of the sentences intact with PBP. The second finding is that improved extended rule-based extraction methods enable multi-word expressions with relation-type components to define new relations without using different linguistic resources. With this extended rule-based structure, multi-word opinion target and opinion words are extracted in the review sentence; thus, this contributes to the extraction resolution performance of OTE. The third finding is that the outputs of the polarity analyzer models are combined by the MV method, and output performance optimization is achieved. Thus, the overall accuracy of the proposed approach has been increased. In addition to all the findings obtained, the process of re-representing the opinion target expressions with words that

Table 4
F-Score results (>0.5) reported in the SemEval16-ABSA Proceeding with abbreviations of the approaches that participated in the Task 5 competition for the OTE dataset [49].

Acronym	Learning	Pre-processing	Model	F-Measure
NLANGP	Supervised	–	RNN+CRF	0.72340
AUEB	Supervised	Checks such as uppercase/lowercase letters, digits, and punctuation marks	CRF	0.70441
Our approach	**Unsupervised**	**Pattern-based**	**Rule-based**	**0.70003**
UWB	Supervised	Lemmatization, POS-N with frequency	CRF	0.67089
GTI	Supervised	POS tagging and lemmatization	CRF	0.66553
SentiSys	Supervised	Review segmentation, sentence tokenizing, sentence tagging	CRF	0.66545
bunji	Supervised	Dictionary creation	RNN	0.64882
DMIS	Supervised	–	–	0.63495
XRCE	Supervised	Tokenization, syntactic parser	CRF+Feedbacked ensemble	0.61980
UWate	Supervised	Seed word extraction	Dep. parsing+MWUs+NPMI	0.57067
KnowC	Supervised	POS tagging	Pre-trained word vec.	0.56816
TGB	Supervised	Removing HTML codes/URLs, tokenization, stemming	Multi-lingual constraint system	0.55054
BUAP	Supervised	–	–	0.50253

Table 5
F-Score results of the proposed approach and other studies associated with the rule-based OTE approaches on the OTE dataset.

Owners	Learning	Pre-processing	Model	F-Measure
Our approach	**Unsupervised**	**Pattern-based**	**Rule-based**	**0.70003**
[10]	Unsupervised	Stopwords removal with WordNet support	Rule-based	0.66870
[5]	Unsupervised	Tokenizing, lowercase	Rule-based	0.64550
[13]	Unsupervised	Word count threshold value and control, and conversion of numeric and numeric+letter words	Rule-based and machine learning	0.64240

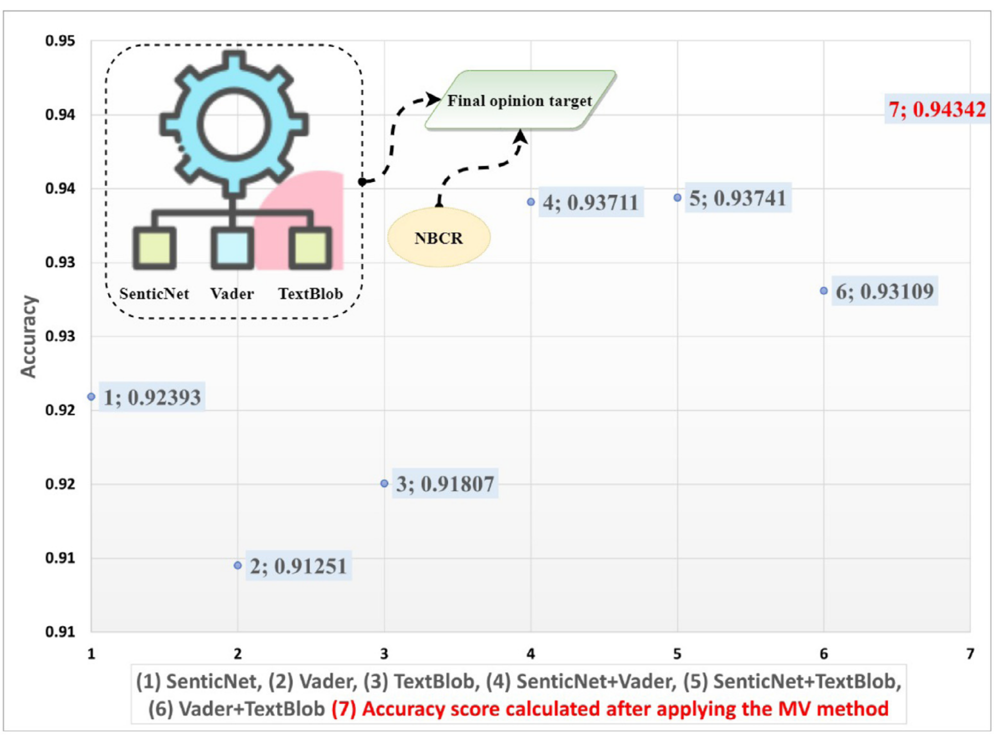


Fig. 7. Calculated accuracy score after the MV method with the measured SenticNet, Vader, TextBlob combinational accuracy scores.

can be used interchangeably are essential to catch the correct expression. For this purpose, the NBCR model was used to detect words with the same reference. This neural model shows a higher performance than the deterministic and statistical models. However, this model runs slower than other models due to its high hardware requirements. On the other hand, coreference resolution models can produce insufficient results on complex sentences. It is provided to use parsed sentences instead of complex sentences in the input of the NBCR model to overcome the problem.

According to the analysis results, it has been determined that the OTE approach obtains effective results for extraction performance. Compared to the related approaches, higher performance was obtained by the proposed approach than other rule-based approaches (Table 5). A comparable F-Measure score was obtained compared with other high-performance approaches in the literature (Table 4).

6. Conclusion

Rule-based techniques with enhanced domain-independent performance as an alternative to more costly supervised methods and unsupervised OTE approach supported by new methods are proposed. In the proposed approach, explicit expressions of the aspects of the entity that are subjectively affected are extracted from the online review texts. The developed PBP, opinion target extraction with extended syntactic-based relation rules, and MV methods are implemented into the proposed approach to improve OTE performance. With the development of the PBP method, a novel text pre-processing was carried out by removing the punctuation marks between sequences of words that do not comply with the determinative group grammar rules in the review texts. With this method, stopwords and other punctuation marks responsible for the meaning and structure of the sentence are preserved. The rule-based techniques using linguistic patterns have been extended with new auxiliary relation types algorithms in extracting opinion targets and opinion words. The MV method is used to optimize the output performance of the proposed approach. In addition, the detection of words with the same reference in the review texts was carried out by the NBCR model, which offers higher performance than other models. In order to improve the resolution performance of this model in complex sentences, the NBCR model input is parsed into simple sentences. In conclusion, the obtained results clearly show that while the proposed OTE approach performs better than other rule-based approaches, it presents results on a comparable scale with other supervised approaches.

From a practical perspective, this study provides businesses with the opportunity to analyze online reviews to improve their products and determine the product aspects that affect the purchasing preference of the consumer. Besides, these analyses allow consumers to detect changes in their product preferences by noticing them. Thus, it will assist managers in decision-making and strategy formulation.

Future work plans to analyze the polarities of the extracted individual opinion terms, the reviews, and the effects of the general polarities in each sentence of the reviews. Moreover, it is aimed to increase the accuracy of the obtained results by improving rule-based aspect extraction methods for English and a few other languages. Consequently, it is aimed to develop domain-independent semantic-based practical and unique approaches on the aspect category extraction.

CRediT authorship contribution statement

Kürşat Mustafa Karaoğlan: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration. **Oğuz Fındık:** Conceptualization, Methodology, Validation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] J. Xliema, Effect of social networks o 11 consumer behaviour: Complex buying, 2019, http://dx.doi.org/10.1016/j.ifacol.2019.12.594.
[2] B.J. Jansen, J.O. Salminen, S.-G. Jung, Making meaningful user segments from datasets using product dissemination and product impact, Data Inf. Manage. 4 (4) (2020) 237–249, http://dx.doi.org/10.2478/dim-2020-0048.
[3] L. Zhang, B. Liu, Aspect and entity extraction for opinion mining, in: Data Mining and Knowledge Discovery for Big Data, Springer, 2014, pp. 1–40, http://dx.doi.org/10.1007/978-3-642-40837-3_1.
[4] B. Ozyurt, M.A. Akcayol, A new topic modeling based approach for aspect extraction in aspect based sentiment analysis: SS-LDA, Expert Syst. Appl. 168 (2021) 114231, http://dx.doi.org/10.1016/j.eswa.2020.114231.
[5] G. Singh Chauhan, Y. Kumar Meena, D. Gopalani, R. Nahta, A two-step hybrid unsupervised model with attention mechanism for aspect extraction, Expert Syst. Appl. 161 (2020) 113673, http://dx.doi.org/10.1016/j.eswa.2020.113673.
[6] K. Ravi, V. Ravi, A survey on opinion mining and sentiment analysis: Tasks, approaches and applications, Knowl.-Based Syst. 89 (2015) 14–46, http://dx.doi.org/10.1016/j.knosys.2015.06.015.
[7] A. García-Pablos, M. Cuadros, G. Rigau, W2VLDA: Almost unsupervised system for aspect based sentiment analysis, Expert Syst. Appl. 91 (2018) 127–137, http://dx.doi.org/10.1016/j.eswa.2017.08.049, arXiv:1705.07687.
[8] T.A. Rana, Y.N. Cheah, Aspect extraction in sentiment analysis: comparative analysis and survey, Artif. Intell. Rev. 46 (4) (2016) 459–483, http://dx.doi.org/10.1007/s10462-016-9472-z.
[9] A.K. J, S. Abirami, Aspect-based opinion ranking framework for product reviews using a spearman's rank correlation coefficient method, Inform. Sci. 460–461 (2018) 23–41, http://dx.doi.org/10.1016/j.ins.2018.05.003.
[10] M. Dragoni, M. Federici, A. Rexha, An unsupervised aspect extraction strategy for monitoring real-time reviews stream, Inf. Process. Manage. 56 (3) (2019) 1103–1118, http://dx.doi.org/10.1016/j.ipm.2018.04.010.
[11] S. Wu, Y. Xu, F. Wu, Z. Yuan, Y. Huang, X. Li, Aspect-based sentiment analysis via fusing multiple sources of textual knowledge, Knowl.-Based Syst. 183 (2019) 104868, http://dx.doi.org/10.1016/j.knosys.2019.104868.
[12] S. Poria, E. Cambria, A. Gelbukh, Aspect extraction for opinion mining with a deep convolutional neural network, Knowl.-Based Syst. 108 (2016) 42–49, http://dx.doi.org/10.1016/j.knosys.2016.06.009.
[13] C. Wu, F. Wu, S. Wu, Z. Yuan, Y. Huang, A hybrid unsupervised method for aspect term and opinion target extraction, Knowl.-Based Syst. 148 (2018) 66–73, http://dx.doi.org/10.1016/j.knosys.2018.01.019.
[14] T.A. Rana, Y.N. Cheah, A two-fold rule-based model for aspect extraction, Expert Syst. Appl. 89 (2017) 273–285, http://dx.doi.org/10.1016/j.eswa.2017.07.047.
[15] F. Hemmatian, M.K. Sohrabi, A survey on classification techniques for opinion mining and sentiment analysis, Artif. Intell. Rev. 52 (3) (2019) 1495–1545, http://dx.doi.org/10.1007/S10462-017-9599-6/FIGURES/10.
[16] S. Sun, C. Luo, J. Chen, A review of natural language processing techniques for opinion mining systems, Inf. Fusion 36 (2017) 10–25, http://dx.doi.org/10.1016/j.inffus.2016.10.004.
[17] H.H. Do, P.W. Prasad, A. Maag, A. Alsadoon, Deep learning for aspect-based sentiment analysis: A comparative review, Expert Syst. Appl. 118 (2019) 272–299, http://dx.doi.org/10.1016/j.eswa.2018.10.003.
[18] H. Gandhi, V. Attar, Extracting aspect terms using CRF and Bi-LSTM models, Procedia Comput. Sci. 167 (2020) 2486–2495, http://dx.doi.org/10.1016/j.procs.2020.03.301.
[19] J. Wen, G. Zhang, H. Zhang, W. Yin, J. Ma, Speculative text mining for document-level sentiment classification, Neurocomputing 412 (2020) 52–62, http://dx.doi.org/10.1016/j.neucom.2020.06.024.
[20] M. Araújo, A. Pereira, F. Benevenuto, A comparative study of machine translation for multilingual sentence-level sentiment analysis, Inform. Sci. 512 (2020) 1078–1102, http://dx.doi.org/10.1016/j.ins.2019.10.031.
[21] B. Liu, Sentiment analysis and opinion mining, Synth. Lect. Hum. Lang. Technol. 5 (1) (2012) 1–184, http://dx.doi.org/10.2200/S00416ED1V01Y201204HLT016.
[22] A.B. Nassif, A. Elnagar, I. Shahin, S. Henno, Deep learning for arabic subjective sentiment analysis: Challenges and research opportunities, Appl. Soft Comput. 98 (2021) 106836, http://dx.doi.org/10.1016/j.asoc.2020.106836.
[23] K. Sailunaz, R. Alhajj, Emotion and sentiment analysis from Twitter text, J. Comput. Sci. 36 (2019) 101003, http://dx.doi.org/10.1016/j.jocs.2019.05.009.
[24] M. Malik, S. Habib, P. Agarwal, A novel approach to web-based review analysis using opinion mining, Procedia Comput. Sci. 132 (2018) 1202–1209, http://dx.doi.org/10.1016/j.procs.2018.05.035.
[25] M. Tubishat, N. Idris, M. Abushariah, Explicit aspects extraction in sentiment analysis using optimal rules combination, Future Gener. Comput. Syst. 114 (2021) 448–480, http://dx.doi.org/10.1016/j.future.2020.08.019.
[26] Y. Wei, H. Zhang, J. Fang, J. Wen, J. Ma, G. Zhang, Joint aspect terms extraction and aspect categories detection via multi-task learning, Expert Syst. Appl. 174 (February) (2021) 114688, http://dx.doi.org/10.1016/j.eswa.2021.114688.

[27] H. Yang, B. Zeng, J. Yang, Y. Song, R. Xu, A multi-task learning model for Chinese-oriented aspect polarity classification and aspect term extraction, Neurocomputing 419 (2021) 344–356, http://dx.doi.org/10.1016/j.neucom.2020.08.001, arXiv:1912.07976.

[28] M. Venugopalan, D. Gupta, An unsupervised hierarchical rule based model for aspect term extraction augmented with pruning strategies, Procedia Comput. Sci. 171 (2020) 22–31, http://dx.doi.org/10.1016/j.procs.2020.04.303.

[29] N. Akhtar, N. Zubair, A. Kumar, T. Ahmad, Aspect based sentiment oriented summarization of hotel reviews, Procedia Comput. Sci. 115 (2017) 563–571, http://dx.doi.org/10.1016/j.procs.2017.09.115.

[30] A.M. Popescu, O. Etzioni, Extracting product features and opinions from reviews, in: HLT/EMNLP 2005 - Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference, 2005, pp. 339–346, http://dx.doi.org/10.3115/1220575.1220618.

[31] C. Sutton, A. McCallum, An introduction to conditional random fields, Found. Trends Mach. Learn. 4 (4) (2011) 267–373, http://dx.doi.org/10.1561/2200000013.

[32] Ł. Augustyniak, T. Kajdanowicz, P. Kazienko, Comprehensive analysis of aspect term extraction methods using various text embeddings, Comput. Speech Lang. 69 (2021) 101217, http://dx.doi.org/10.1016/j.csl.2021.101217.

[33] T. Wang, Y. Cai, H.F. Leung, R.Y. Lau, Q. Li, H. Min, Product aspect extraction supervised with online domain knowledge, Knowl.-Based Syst. 71 (2014) 86–100, http://dx.doi.org/10.1016/j.knosys.2014.05.018.

[34] W. Song, Z. Wen, Z. Xiao, S.C. Park, Semantics perception and refinement network for aspect-based sentiment analysis, Knowl.-Based Syst. 214 (2021) 106755, http://dx.doi.org/10.1016/j.knosys.2021.106755.

[35] D. Anand, B.S. Mampilli, A novel evolutionary approach for learning syntactic features for cross domain opinion target extraction, Appl. Soft Comput. 102 (2021) 107086, http://dx.doi.org/10.1016/j.asoc.2021.107086.

[36] S. Li, L. Zhou, Y. Li, Improving aspect extraction by augmenting a frequency-based method with web-based similarity measures, Inf. Process. Manage. 51 (1) (2015) 58–67, http://dx.doi.org/10.1016/j.ipm.2014.08.005.

[37] A. Moreno-Ortiz, J. Fernández-Cruz, Identifying polarity in financial texts for sentiment analysis: A corpus-based approach, Procedia - Soc. Behav. Sci. 198 (2015) 330–338, http://dx.doi.org/10.1016/j.sbspro.2015.07.451.

[38] A.S. Nayak, A.P. Kanive, Survey on pre-processing techniques for text mining, Int. J. Eng. Comput. Sci. 5 (6) (2016) 16875–16879, http://dx.doi.org/10.18535/ijecs/v5i6.25.

[39] Q. Liu, Z. Gao, B. Liu, Y. Zhang, Automated rule selection for opinion target extraction, Knowl.-Based Syst. 104 (2016) 74–88, http://dx.doi.org/10.1016/j.knosys.2016.04.010.

[40] G. Qiu, B. Liu, J. Bu, C. Chen, Opinion word expansion and target extraction through double propagation, Comput. Linguist. 37 (1) (2011) 9–27, http://dx.doi.org/10.1162/coli_a_00034.

[41] D. Ruta, B. Gabrys, Classifier selection for majority voting, Inf. Fusion 6 (1) (2005) 63–81, http://dx.doi.org/10.1016/j.inffus.2004.04.008.

[42] X. Dong, Z. Yu, W. Cao, Y. Shi, Q. Ma, A survey on ensemble learning, Front. Comput. Sci. 14 (2) (2020) 241–258, http://dx.doi.org/10.1007/s11704-019-8208-z.

[43] L.I. Kuncheva, C.J. Whitaker, C.A. Shipp, R.P. Duin, Limits on the majority vote accuracy in classifier fusion, Pattern Anal. Appl. 6 (1) (2003) 22–31, http://dx.doi.org/10.1007/s10044-002-0173-7.

[44] S.S. Sharma, G. Dutta, SentiDraw: Using star ratings of reviews to develop domain specific sentiment lexicon for polarity determination, Inf. Process. Manage. 58 (1) (2021) 102412, http://dx.doi.org/10.1016/j.ipm.2020.102412.

[45] B. Liang, H. Su, L. Gui, E. Cambria, R. Xu, Aspect-based sentiment analysis via affective knowledge enhanced graph convolutional networks, Knowl.-Based Syst. 235 (2022) 107643, http://dx.doi.org/10.1016/j.knosys.2021.107643.

[46] H. Gauba, P. Kumar, P.P. Roy, P. Singh, D.P. Dogra, B. Raman, Prediction of advertisement preference by fusing EEG response and sentiment analysis, Neural Netw. 92 (2017) 77–88, http://dx.doi.org/10.1016/j.neunet.2017.01.013.

[47] A. Borg, M. Boldt, Using VADER sentiment and SVM for predicting customer response sentiment, Expert Syst. Appl. 162 (2020) 113746, http://dx.doi.org/10.1016/j.eswa.2020.113746.

[48] C.J. Hutto, E. Gilbert, VADER: A parsimonious rule-based model for sentiment analysis of social media text, in: Proceedings of the 8th International Conference on Weblogs and Social Media, in: ICWSM 2014, vol. 8, 2014, pp. 216–225.

[49] M. Pontiki, D. Galanis, H. Papageorgiou, I. Androutsopoulos, S. Manandhar, M. Al-Smadi, M. Al-Ayyoub, Y. Zhao, B. Qin, O. De Clercq, et al., Semeval-2016 task 5: Aspect based sentiment analysis, in: International Workshop on Semantic Evaluation, 2016, pp. 19–30.