Aspect Term Extraction Using Different Methods

Alok Kumar, Department of Computer Science and Engineering, University Institute of Engineering and Technology, Chhatrapati Shahu Ji Maharaj University, Kanpur-208024, India, akumar.uiet@gmail.com
Ruchika Rastogi, Department of Management, Pranveer Singh Institute of Technology, Kanpur 209305, India
rastogi.ruchika@gmail.com
Deepak Kumar Verma , Department of Computer Science and Engineering, University Institute of Engineering and Technology, Chhatrapati Shahu Ji Maharaj University, Kanpur-208024, India. deepak300572@gmail.com

ABSTRACT

The "Aspect Based Sentiment Analysis" (ABSA) task is a classic Sentiment Analysis problem that focuses on understanding people's opinion on aspect level. This main task is sub-divided into multiple works including the extraction of aspect term, recognition of aspect category, recognition of sentiment word and classification of emotions (positive, negative, conflict, neutral) for aspect term and also for aspect category. Here we propose the system for recognizing aspects and analyzing the sentiments using different methods for the laptop review dataset from SemEval-2014. Sentiment Analysis is the process of computationally identifying and categorizing opinions expressed in a piece of text, especially in order to determine the writer's attitude towards a particular topic or product.

INTRODUCTION

Initial Sentiment analysis tasks performed are simply make the prediction on the sentence level or the document level recognizing the sentiment on the sentence or the document level. Through this process a single sentiment is picked out from the entire sentence. During the time more precise analysis of opinions and sentiments regarding the characteristics mentioned in the sentence is required. This pull attention towards "Aspect Based Sentiment Analysis" (ABSA). In ABSA the task it is considered that the opinion of the writer in the entire sentence or document is subject to change with respect to different characteristics of the entity mentioned. Entity can be product of E-commerce like in our case laptop and it's characteristics are colour, size, price, screen etc.

Sentiment analysis is defined as the task of finding the opinions of authors about specific entities. In making decisions, people take into account the opinions of both thought leaders and ordinary people. When a person wants to buy a product online he or she will typically start by searching for reviews and opinions written by other people on the various offerings. Sentiment analysis is one of the hottest research areas in computer science. A sentiment analysis model is used to analyze a text string and classify it with one of the labels that you provide; for example, you could analyze a tweet to determine whether it is positive or negative, or analyze an email to determine whether it is happy, frustrated, or sad. Aspect-based sentiment analysis is the research problem that focuses on the recognition of all sentiment expressions within a given document and the aspects to which they refer. It is common to classify sentences into two principal classes with regard to subjectivity: objective sentences that contain factual information and subjective sentences that contain explicit opinions, beliefs, and views about specific entities. Here, I mostly focus on analyzing subjective sentences. As an example, here is a review about a hotel in Manhattan. "The king suite was spacious, clean, and well appointed. The reception staff, bellmen, and housekeeping were very helpful. Requests for extras from the maid were always provided. The heating and air conditioning functioned well; this was good as the weather was variable. The sofa bed was the best I've ever experienced. The king size bed was very comfortable. The building and rooms are very well soundproofed. The neighborhood is the best for shopping, restaurants, and access to the subway. Only "complaint" has to do with high-speed Internet access. It's only available on floors 8–12."Overall the review is very positive about the hotel. It refers to many different aspects of the hotel including: heating, air conditioning, staff courtesy, bed, neighborhood, and Internet access. In this paper we propose sentiment analysis systems which are able to provide a sentiment score for the whole review as well as analyze the sentiment of each individual aspect of the hotel. In this paper we propose an aspect based

sentiment analysis task of SemEval 2014
[6] which deals with the laptop review dataset.
Particularly we have considered the two subtasks :

- Identification of aspect term which is usually present in the sentence
- Finally sentiment polarity detection for the aspect term.

RELATED WORK

The purpose of sentiment analysis is to analyze the attitudes and excitements of individuals regarding certain entities and the expressions, viewpoints and opinions that they have regarding those entities based on computational analysis. There are five components that make up the opinion sentiment, namely e,t,s,h,dt, a quintuple of the five components of the opinion sentiment (e,t,s,h,dt), where e represents the entity and t represents the target aspect or feature for the sentiment, s represents the sentiment towards the target, h represents the opinion holder, and dt represents the time stamp of the sentiment.
Sentiment analysis can be conducted at three levels: the document level, the sentence level, and the aspect level. At the document level, the whole review is considered a unit of basic information and then classified as either positive, negative, or neutral based on its general sentiment. In the same way, at the sentence level, all of the sentences are regarded as short documents, and the sentiments of all these sentences are identified. The sentiment at the document level is determined based on the entire document as a whole, and the sentiment associated with the target is clearly defined at this level. In this case, the opinion holder might have a positive opinion regarding the entity but might not be satisfied with all the "aspects" of that target. To extract such information, aspect-based classification is used. The aspect-based approach covers both the entity and the aspect aspects in the analysis. It does so by decomposing the entity into aspects (aspect extraction), then classify each aspect sentiment into positive, negative or neutral (aspect sentiment classification), and finally, summarize the results of the previous steps.

Another sentiment analysis level that has been considered is the concept-level [B]. Unlike word-based approaches, concept-level sentiment analysis focuses on the semantic analysis of text through the use of web ontologies and semantic networks. This allows the combination of conceptual and affective words associated with natural language. For example, if the concept "Cloud Computing" is split into two words, the word "cloud" would be wrongly associated with the weather. However, concept-level sentiment analysis is limited by the bounds of the knowledge base and by the fact that it fails to detect important discourse structure information that is essential for effectively detecting the polarity expressed by natural language opinions [C].
In this paper, our concern is aspect-based sentiment analysis, which is one of the levels of sentiment classification. It comprises aspect or feature extraction, sentiment polarity prediction. Aspects extraction in sentiment analysis is now becoming an active area of research as it is the most vital task in the aspect-based recognition [A]. Aspect-based sentiment analysis is the process in which sentiments in respect to different aspects are detected [F]. Aspects are attributes, characteristics, or features of a product or service. Aspect extraction phase involves the identification of these review characteristics through consumers' comments to identify aspects. Next, the polarity prediction and classification take place to decide if the aspect sentiment polarity denotes positive, negative or neutral orientation as well as its strength or tone level [G]. The last step is to summarize the results according to the extracted aspects and their corresponding classified polarity. This summary is essential to determine the strengths and weaknesses of each aspect within the app and compared to others. This summarization of results can be done, qualitatively through text-based aggregated opinions summary [H], or quantitatively through graphical and analytical representation.
To automate the summarization of reviews, aspect words are grouped into aspects categories. In an

unsupervised machine learning approach, the model tries to make sense of the data and extract features on its own. However, when the aspect categories are known in advance, and there is enough training data available, a supervised machine learning approach to aspect category and polarity detection is feasible and may yield better results [I]. On the other hand, semi-supervised approaches use a small set of labeled data to label a larger set of unlabeled data [J].
Reviews aspects are domain dependent and differ from a context to another. Aspect-based sentiment analysis has been broadly used in numerous application domains such as products reviews, social media, hotel reviews and restaurant reviews.
Researchers have reported various approaches in order to extract aspects from textual resources. For instance, Samha et al. [16] used frequent Part of Speech (POS) tags and rules in addition to opinion lexicon to identify aspects and opinion words from reviews as well as group them into categories and summarize the results. POS is the process of marking up a word in a text as corresponding to a particular part of speech (i.e. verb, adjective, preposition, etc.), according to the word's definition and context [K]. Devi et al. [L]proposed a feature-based approach for sentiment analysis using Support Vector Machine (SVM). The authors collected product reviews on laptops from e-commerce platforms such as Amazon and eBay. SentiWordNet, which is a lexical resource for sentiment classification and opinion mining applications, was used to identify objective sentences and later to identify the polarity for the opinion words. Furthermore, POS tagging was used to extract aspect terms from the dataset. The authors used Stanford parser to extract the opinion words and to find the grammatical dependencies to determine the connection between the opinion words and aspects extracted in the previous step. The dependencies also assist in determining the negations that were considered in calculating the polarity score. The SVM classifier was used to classify aspects and determine their sentiment polarity score. The result of the SVM classifier is a set of vectors that contain aspects and its opinion words for each review. The performance achieved significant results with overall accuracy of 88.16%.
Similarly, Manek et al. [M] proposed a feature selection method based on Gini index. The authors used SVM classifier to predict sentiment polarities for a movies' reviews dataset. In their approach, the reviews were pre-processed with tokenization, case transformation, filtering stop words and stemming. Further, the Term Frequency/Inverted Document Frequency (TF/IDF) has been adopted with the weighting mechanism using Gini Index as a feature selection approach. This helped in measuring the impurity of the attribute for categorization and creation of the feature vector for the top 50 attributes according to the Gini impurity index value. Then, the SVM classifier was applied in order to train and test the model. This approach achieved significant results with overall accuracy of 92.81%.The introduction of the SemEval competition resulted in a rise to the number of proposed methods for aspects extractions. For instance, Mubarok et al. [N] used sentiment analysis and classification techniques to determine the sentiment polarity of restaurants reviews using the SemEval 2014 dataset. Feature extraction was performed using Chi Square, resulting in higher computational speed despite reducing the system performance. Naïve Bayes classification of sentiment polarity was used to classify both aspects and sentiments. The evaluation results indicated that the system performed well with a highest score of 78.12% for the F1-Measure.
On the other hand, García-Pablos et al. [O] presented W2VLDA, an aspect sentiment classification system that requires minimal supervision and does not require language or domain specific resources. The system is able to distinguish opinion words from aspect terms in an unsupervised way. The only supervision required by the user is a single seed word per aspect and polarity. The system performance was also evaluated using the SemEval 2016 dataset. The analysis showed competitive results for different languages and domains. Similarly, Dragoni et al. [P] presented an aspect-based unsupervised system for opinion monitoring that supports data visualization. The authors adopted an open information extraction approach to extract the aspects. The system aimed

at providing users with an effective analysis and visualization tool based on the user-generated content. The approach proved its effectiveness compared to baseline supervised approaches participated in SemEval campaign.

On the other hand, Xianghua et al. [Q] suggested an unsupervised approach to determine the aspects and sentiments in Chinese social reviews. The authors used Latent Dirichlet Allotiment (LDA) in social review to identify multi-aspect global topics. Then, they extracted the local topic and the related sentiment according to a sliding window through the review text. Chen et al. [R] improved the LDA model and presented the Automated Knowledge LDA which is a fully automatic approach that can use existing domain independent data to learn prior knowledge and identify new aspects. The Automated Knowledge LDA approach was able to produce aspects and resolve issues related to wrong knowledge by adopting and enhancing the Gibbs sampler method. Moreover, Poria et al. [S] asserted an unsupervised rule-based approach to obtain both explicit aspects as well as implicit aspects clues (IAC) from product and restaurant reviews. In the proposed approach, the authors used the IACs to identify the IACs in the review and plot them to the aspects they are signifying depending on common-sense knowledge and on the dependency structure of the sentence using WordNet and SenticNet.

Liu et al. [T] integrated both supervised and unsupervised domain independent automatic rule-based methods to improve double propagation. In their research, the double propagation method assumes that the opinion words will always have a target. Therefore, there is a syntactic relation between the opinion word and the target within the same sentence.

As can be noticed, most of the aforementioned works require labeled datasets for training their models for each of the domains. It would be difficult to train the models when studying a new domain without first extracting the domain-specific characteristics. It will be possible for the models to capture the aspect specific sentiment with minimal data requirements when domain-independent parameters are used to describe the relationships between aspects and their associated opinion expressions.

TASK DEFINITION

Using natural language processing, text analysis, and computational linguistics, sentiment analysis (also known as opinion mining) identifies and extracts subjective information from sources. The main function of sentiment analysis is to determine the polarity of the document, sentence, entities/aspect whether the opinion expressed is positive, negative or neutral. There are four main task under sentiment analysis namely, aspect term extraction, aspect term polarity, aspect category detection, and aspect category polarity. In contrast, here we focuses on aspect based sentiment analysis where the goal is to identify aspects of the given target entity and sentiment expressed towards each aspect. In particular, the paper focuses on two tasks of SemEval 2014 Conference.

A. Subtask 1:

Aspect Term Extraction: The task is to identify all the distinct aspect terms present in a set of sentences containing predefined entities (e.g. laptops) and return a list that contains all the distinct aspect terms from the set of sentences.

B. Subtask 2:

Aspect Term Polarity: An aspect term set within a sentence must be analyzed in order to determine whether it falls within one or more polarities, i.e., whether it is either positive, negative, neutral, or contradictory with other aspects.

C. Laptop Reviews Dataset

This dataset provided as a trial dataset in SemEval 2014 Conference, consists of over 3K English sentences from the laptop reviews. The sentences in the datasets are annotated using XML tags. It is shown in the following example how the annotated sentences from the laptop dataset are formatted.

XML Format

```
<sentence id=”533”>
    <text>
        From the build quality to the performance, everything about it has been sub-par from
        what I would have expected from Apple.
    </text>
    <aspectTerms>
        <aspectTerm term=”build quality”polarity=”negative” from=”9”
            to=”22”/>
        <aspectTerm term=”performance”polarity=”negative” from=”30”
            to=”41”/>
    </aspectTerms>
</sentence>
```

In the sentences of this dataset, there is an <aspectTerm … /> element for each instance of an aspect term. For example, if the previous sentence contained two occurrences of the aspect term “performance”, there would be two <aspectTerm …/> elements, which would be identical if both occurrences had negative polarity. If a sentence has no aspect terms, there is no <aspectTerms> …</aspectTerms> element in its annotation.

Note that the sentences may contain spelling mistakes. The identified aspect terms should be returned as they appear in the sentences, even if misspelled (e.g., “warranty” as “warrenty”).

For each aspect term of the training data we include two attributes (“from” and “to”) that indicate its start and end offset in the text (e.g., <aspectTerm term=”staff” polarity=”negative” from=”8” to=”13”>).

The possible values of the polarity field are: positive, negative‖, conflict, ―neutral

ASPECT BASED SENTIMENT ANALYSIS

In this task, aspect terms will be identified from the sentence, and correspondingly, aspect categories will be recognised for the laptop review.

MAIN FEATURES

The process of determining the features of a text is one of the most important phases in the process of extracting information from it. The features are the elements algorithms use as input data for training and classification. The main features to be used in the development in this work are presented as follows:

1. **Tokenization**, as defined by Manning et al. (2014), is the task of chopping it up into pieces, called tokens, perhaps at the same time throwing away certain characters, such as punctuation. These tokens are often loosely referred to as terms or words, but it is sometimes important to make a type/token distinction. There is a term token used in semantics to refer to the sequence of characters within a document that are grouped together as a useful semantic unit for processing such a document. A type is the class of all tokens containing the same character sequence.

2. **Stemming** is done by removing any attached suffixes and prefixes (affixes) from index terms before the actual assignment of the term to the index (JIVANI et al., 2011). As a part of the linguistic study of morphology, stemming is also part of the field of artificial intelligence (AI), the retrieval and extraction of information. With stemming and artificial intelligence, relevant information can be extracted from vast sources such as big data or the Internet. For example, additional forms of a word may need to be searched to achieve the best results.. Stemming is too a part of queries and Internet search engines.

3.Lemmatization is the process of finding an word that can represent every form of a word. In this process, words are analyzed using a vocabulary and morphological analysis, normally aiming to remove inflectional endings only and return to their base form, known as lemma..

3. **PoS Tagging** is the method of correlation between a word in a sentence and its Part of Speech. It is a method based both on the definition of the word as well as its relationship to other words, which give it context.

4. **Chunking** is the hierarchy of ideas in a text, it gives the ability to the speaker to generalize or specify a word inside a sentence. When the speaker churns up or down, it allows the speaker to use certain language patterns, to utilize the natural internal process of language, or to reach for more specific information, bits, or portions that are missing from the speech.

Aspect Term Extraction

An entity of the target is defined by its particular aspect term. The term is defined by its unique position in a text. It may not be explicit, and can be expressed by pronouns or text coreferences. For explicit target extraction, there are three main approaches:

1. Noun-based extraction: Initially developed by Hu and Liu (2004), uses a grammar analyzer to identify the most frequent nouns.
2. Sentiment and target relation extraction: Uses a grammar analyzer and dependency relations to find relation between sentiment words and their targets.
3. Supervised learning extraction: Uses supervised machine learning models to determine if an opinion is about an entity or an aspect of an entity.

Aspect Term Polarity

Each opinion in a sentence has a polarity from the set P = {positive, negative, neutral}. There is a situation when there is no clear definition of polarity when it comes to a classification of neutral sentiment. Two main approaches are used to determine an opinion's polarity:

1. **Supervised learning based attribution**: In this approach, supervised learning is used to determine a sentence's opinion at the sentence level. The sentence can be the scope of the sentiment expression. This approach makes the method dependent on the training data, yielding poorer results when applied to differentdomains (LIU, 2012).
2. **Lexical information based attribution**: As a set of methods that are usually supervised, they rely on opinion dictionaries and other processing resources such as grammar analyzers or dependency trees in order to determine the degree of polarity of an opinion.

METHODS

1. USING STANZA OR DEPENDENCYPARSER

Stanza is a collection of accurate and efficient tools for the linguistic analysis of many human languages. Starting from raw text to syntactic analysis and entity recognition, Stanza brings state-of-the-art NLP models to languages of your choosing.

The term Dependency Parsing (DP) refers to the process of examining the dependencies between the phrases of a sentence in order to determine its grammatical structure. A sentenceis divided into many sections based mostly on this. The process is based on the assumption that there is a direct relationship between each linguistic unit in a sentence. These hyperlinks are called dependencies.

To extract aspect terms from review using stanza and dependency parser we need to make different rules that should follow to extract aspect and terms related to it to detect its polarity.

2. USING NAME ENTITY RECOGNITION

The named entity recognition (NER) is one of the most data preprocessing tasks. It involves the identification of key information in the text and classification into a set of predefined categories. An entity is basically the thing that isconsistently talked about or referred to in the text.

NER is the form of NLP.

NER model is also used for aspect based sentiment analysis. For aspect term extractionwe take a blank NER model and train it according to our data and get result for new review.

We need to prepare data in the given format:

[('I charge it at night and skip taking the cord with me because of the good battery life.',{'entities': [(41, 45, 'ASPECT')]}),('The techguy then said the service center does not do

1-to-1 exchange and I have to direct my concern to the "sales" team, which is the retail shop which I bought my netbook from.',{'entities': [(27, 41, 'ASPECT')]})]

We need to create a list of a tuple. A tuple contain review and dictionary of entities related to review(from,to,label).

3. USING MACHINE LEARNINGMETHODS

All machine learning models are categorized as either **supervised** or **unsupervised**. If the model is a supervised model, it's then sub-categorized as either a **regression** or **classification** model.

Supervised learning involves learning a functionthat maps an input to an output.

NAIVE BAYES CLASSIFIERS

Naive Bayes classifiers are a collection of classification algorithms based on Bayes' Theorem. It is not a single algorithm but a family of algorithms where all of them share a common principle, i.e. every pair of features being classified is independent of each other.

SUPPORT VECTOR MACHINE

A **Support Vector Machine** is a supervised classification technique that can actually get pretty complicated but is pretty intuitive at themost fundamental level.

Let's assume that there are two classes of data. A support vector machine will find a **hyperplane** or a boundary between the two classes of data that maximizes the margin between the two classes (see below). There are many planes that can separate the two classes, but only one plane can maximize the margin or distance between the classes.

STOCHASTIC GRADIENTDESCENT (SGD)

SGD is a simple yet efficient optimization algorithm used to find the values of parameters/coefficients of functions that minimize a cost function. In other words, it is used for discriminative learning of linear classifiers under convex loss functions such as SVM and Logistic Regression.

DECISION TREE

Decision trees are a popular model, used in operations research, strategic planning, andmachine learning. Each square above is called a **node**, and the more nodes you have, the more accurate your decision tree will be (generally). The last nodes of the decision tree, where a decision is made, are called the **leaves** of the tree. Decision trees are intuitive and easy to build but fall short when it comes to accuracy.

RANDOM FOREST

Random forests are an ensemble learning technique that builds off of decision trees. Random forests involve creating multiple decision trees using bootstrapped datasets of the original data and randomly selecting a subset of variables at each step of the decision tree. The model then selects the mode of all of the predictions of each decision tree.

Data Preparation for Machine Learning Methods

1. Read XML file and separate review andits aspect information(aspect info means aspect term, from, to, polarity)
2. Replace all negation from the review. Ex. "The camera quality is not good." Remove not for the review, then it becomes "The camera quality is bad."
3. Clean the review by removing stopwords and punctuation marks from review. Also perform stemming or lemmatization, POS tagging etc to cleanit.
4. Count the frequency of all aspect terms given to us and take only 1000 aspect terms as output of the model.
5. Now all the aspect terms become columns and all reviews become rows. Fill a row-column pair with 0 or 1 i.e. whether the aspect term is present in the given review or not.
6. As input cannot be sent as string soconvert this into a vector using CountVectorizer.
7. Now split the dataset into train data and test data (75% train data and 25%test data).
8. Fit the model using train data and testthe results for test data.

PERFORMANCE EVALUATION

Accuracy performance metrics can be decisive when dealing with imbalanced data.The confusion matrix, precision, recall, and F1 score gives better intuition of prediction results as compared to accuracy.

What is a confusion matrix?

It is a matrix of size 2×2 for binary classification with actual values on one axis and predicted on another.

True Positive (TP) — model correctly predicts the positive class (prediction and actual both are positive).

True Negative (TN) — model correctly predicts the negative class (prediction and actual both are negative).

False Positive (FP) — model gives the wrong prediction of the negative class (predicted-positive, actual-negative).

False Negative (FN) — model wrongly predicts the positive class (predicted-negative, actual-positive).

Precision

Out of the entire positive predicted what percentage is truly positive.

The precision value lies between 0 and 1.

Recall

Out of the total positive, what percentage are predicted positive. It is the same as TPR (true

$$Recall = \frac{TP}{TP + FN}$$

positive rate).

F1 Score

It is the harmonic mean of precision and recall. It takes both false positive and false negatives into account. Therefore, it performs well on an imbalanced dataset.

$$F1\ score = \frac{2}{\frac{1}{Precision} + \frac{1}{Recall}} = \frac{2 * (Precision * Recall)}{(Precision + Recall)}$$

F1 score gives the same weightage to recall and precision. There is a **weighted F1 score** in whichwe can give different weightage to recall and precision.

CONCLUSIONS

Aspect-based sentiment analysis is considered as one of the challenging tasks in sentiment analysis area of research. It is important that all feedbacks are understood and categorized so that smart government can rely on this channelto listen to their customers. Therefore, this canbe considered as a factor for future smart services improvements and optimizations that exceed the people's expectations.

REFERENCES

A. B.Liu, Sentiment analysis and opinion mining, Synthesis Lectures Human Lang. Technol. 5 (2012) 1–167, http://dx.doi.org/10.2200/S00416ED1V01Y201204HLT016.

B. S. Poria, E. Cambria, G. Winterstein, G.-B. Huang, Sentic patterns: Dependency-based rules for concept-level sentiment analysis, Knowl.-Based Syst. 69 (2014) 45–63, http://dx.doi.org/10.1016/j.knosys.2014.05.005.

C. S. Poria, E. Cambria, A. Gelbukh, F. Bisio, Hussain, Sentiment data flow analysis by means of dynamic linguistic patterns, IEEE Comput. Intell. Mag. 10 (2015) 26–36, http://dx.doi.org/10.1109/MCI.2015.2471215.

D. M.Z. Asghar, A. Khan, S. Ahmad, F.M. Kundi, A review of feature extraction insentiment analysis, J. Basic Appl. Scientific Res. 4 (2014) 181–186.

E. S.Y. Ganeshbhai, Feature based opinion mining : a survey, in: Advance Computing Conference (IACC), 2015 IEEE International, IEEE, 2015, pp. 919–923.

F. M.S. Akhtar, D. Gupta, A. Ekbal, P. Bhattacharyya, Feature selection and ensemble construction: a two-step method for aspect based sentiment analysis, Knowl.-Based Syst. 125 (2017) 116–135, http://dx.doi.org/10.1016/j.knosys.2017.03.020.

G. E. Cambria, Affective computing and sentiment analysis, IEEE Intell. Syst. 31 (2016) 102–107, http://dx.doi.org/10.1109/MIS.2016.31.

H. G. Carenini, J.C.K. Cheung, A. Pauls, Multi-document summarization of evaluative text, Comput. Intell. 29 (2013) 305–312.

I. K. Schouten, O. van der Weijde, F. Frasincar, R. Dekker, Supervised and unsupervised aspect category detection for sentiment analysis with co-occurrence data, IEEE Trans. Cybern. 48 (2018) 1263–1275,http://dx.doi.org/10.1109/TCYB.2017.2688801.

J. B. Liu, L. Zhang, A survey of opinion mining and sentiment analysis, in: Mining Text Data, Springer US, 2012, pp.415–463.

K. A.K. Samha, Y. Li, J. Zhang, Aspect-based opinion extraction from customer reviews, in: Computer Science and Information Technology (CS and IT), Volume 4, Number 4: Proceedings of the Second International Conference of Database and Data Mining (DBDM 2014), AIRCC Publishing Corporation, 2014, pp. 149–160.

L. D.V.N. Devi, C.K. Kumar, S. Prasad, A feature based approach for sentiment analysis by using support vector machine, in: 2016 IEEE 6th International Conference on Advanced Computing (IACC), IEEE, 2016, pp. 3–8, http://dx.doi.org/10.1109/IACC.2016.11

M. A.S. Manek, P.D. Shenoy, M.C. Mohan, K.R. Venugopal, Aspect term extraction for sentiment analysis in large movie reviews using Gini Index feature selection method and SVM classifier, World Wide Web. 20 (2017) 135–154, http://dx.doi.org/10.1007/s11280-015-0381-x.

N. M.S. Mubarok, M. Adiwijiaya, Aldhi Dwi, Aspect-based sentiment analysis to review products using Naïve bayes, in: AIP Conference Proceedings, AIP Publishing, 2017, pp. 1–8, http://dx.doi.org/10.1063/1.4994463.

O. García-Pablos, M. Cuadros, G. Rigau, W2VLDA: almost unsupervised system for aspect based sentiment analysis, Expert Syst. Appl. 91 (2018) 127–137.

P. M. Dragoni, M. Federici, A. Rexha, An unsupervised aspect extraction strategy for monitoring real-time reviews stream, Inf. Process. Manage. (2018) 1103–1118.

Q. Xianghua, L. Guo, G. Yanyan, W. Zhiqiang, Multi-aspect sentiment analysis for Chinese online social reviews based on topic modeling and HowNet lexicon, Knowl.-Based Syst. 37 (2013) 186–195, http://dx.doi.org/10.1016/j.knosys.2012.08.003.

R. Z. Chen, A. Mukherjee, B. Liu, Aspect Extraction with Automated Prior Knowledge Learning, in: Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2014: pp. 347–358.

S. S. Poria, E. Cambria, L.-W. Ku, C. Gui, A. Gelbukh, A rule-based approach to aspect extraction from product reviews, in: Proceedings of the Second Workshop on Natural Language Processing for Social Media (SocialNLP), 2014, pp. 28–37.

T. Q. Liu, Z. Gao, B. Liu, Y. Zhang, Automated rule selection for aspect extraction in opinion mining, in: IJCAI,2015: pp. 1291–1297.

U. S. Mukherjee, P. Bhattacharyya, Feature specific sentiment analysis for product reviews, in: A. Gelbukh (Ed.), Computational Linguistics and Intelligent Text Processing, Springer, Berlin Heidelberg, 2012, pp. 475–487.

V. Vilares, C. Gómez-Rodríguez, M.A. Alonso, Universal, unsupervised (rule-based), uncovered sentiment analysis, Knowl.-Based Syst. 118 (2017) 45–55.

W. M. Taboada, B. Julian, T. Milan, V. Kimberly, S. Manfred, Lexicon-based methods for sentiment analysis, Comput. Linguistics 37 (2011) 267–307.

X. F.M. Kundi, A. Khan, S. Ahmad, M.Z. Asghar, Lexicon-based sentiment analysis in the social web, J. Basic Appl.Scientific Res. 4 (2014) 238–248.