IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 23, 2017, accepted January 19, 2018, date of publication January 24, 2018, date of current version February 28, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2797224

Opinion–Aspect Relations in Cognizing Customer Feelings via Reviews

ANH-DUNG VO, QUANG-PHUOC NGUYEN, AND CHEOL-YOUNG OCK
Department of IT Convergence, University of Ulsan, Ulsan 44610, South Korea

Corresponding author: Cheol-Young Ock (okcy@ulsan.ac.kr)

This work was supported by the ICT R&D Program of MSIP/IITP(Development of Core Technology for Context -aware Deep-Symbolic Hybrid Learning and Construction of Language Resources) under Grant 2013-0-00179.

ABSTRACT Determining a consensus opinion on a product sold online is no longer easy, because assessments have become more and more numerous on the Internet. To address this problem, researchers have used various approaches, such as looking for feelings expressed in the documents and exploring the appearance and syntax of reviews. Aspect-based evaluation is the most important aspect of opinion mining, and researchers are becoming more interested in product aspect extraction; however, more complex algorithms are needed to address this issue precisely with large data sets. This paper introduces a method to extract and summarize product aspects and corresponding opinions from a large number of product reviews in a specific domain. We maximize the accuracy and usefulness of the review summaries by leveraging knowledge about product aspect extraction and providing both an appropriate level of detail and rich representation capabilities. The results show that the proposed system achieves F1-scores of 0.714 for camera reviews and 0.774 for laptop reviews.

INDEX TERMS Aspect-based, expert system, knowledge acquisition, sentiment analysis, text mining.

I. INTRODUCTION

With the rapid growth of e-commerce, millions of customers can now share their opinions about many kinds of products in discussion groups, merchant sites, their personal blog, or a review website. As a result, the number of online product reviews available on the internet is increasing rapidly. Already, the number of available reviews makes it impractical for prospective customers to read them all and discern a consensus opinion about a product. Therefore, automated opinion detection and summarization systems have emerged to help people make an informed decision

Opinion mining, also known as sentiment analysis, has grown out of this need. Sentiment analysis extracts valuable subjective information from the raw text of reviews. Opinion mining can be divided into three tasks: sentiment classification (document-level), subjective/objective identification (sentence-level), and aspect-based sentiment analysis (feature level). Sentiment classification is the most broadly researched topic; it classifies a review as positive or negative. Subjective/objective identification identifies subjective sentences that can include sentiments. However, both the document-level and sentence-level classifications are too coarse for most current applications because they cannot determine exactly what people liked. Summarizing the subjective information in customer reviews based on related aspects and sentiments is a more effective way to help customers efficiently digest the enormous amount of available information.

In this paper, we propose a system to extract product aspects and corresponding opinions from online product reviews. As our main contribution, we introduce a system that includes two stages—knowledge extraction and sentiment analysis. In the first stage, the system takes a two-step approach to extract syntactic knowledge and implied opinion–aspect relations using a set of natural language processing (NLP) tools. First, coarse knowledge from reviews is automatically obtained using dependency parser (DP), co-reference (CR), and named entity recognition (NER) tools. Second, additional opinion–aspect relations are inferred from aggregate statistics for the extracted coarse knowledge. In the second stage, the knowledge from the first stage is used to analyze new reviews and generate aspect-based summaries. No NLP systems are perfect, but this system improves the accuracy and usefulness of its review summaries by leveraging knowledge about product aspect extraction and providing both an appropriate level of detail and rich representation capabilities. Extensions of this work are described in the discussion session.

II. LITERATURE REVIEW

A. TEXT SUMMARIZATION

Text summarization [1]–[4] focuses on identifying and extracting the main entities and facts from a raw text document. The extraction framework detects discrete portions of the text that are most representative of the document's content. Most existing experiments on text summarization focus on an individual document. Recently, researchers [5] have considered text summarization of multiple documents about related information. The summaries are generated by selecting sentences that address the most specific word associations within the documents. Those approaches rely on the strength of word associations in the set of documents to be summarized.

Our work, aspect-based summaries of customer reviews, is related to but different from ordinary text summarization in several ways. First, we do not summarize documents by picking or rewording a subset of the original sentences to obtain their main information, as in common text summarization. Instead, we identify and extract certain product aspects and the corresponding opinions about them from online reviews. Second, we do not focus on facts; we observe and extract subjective information and opinions based on facts. Third, a summary in our system is structured using opinion–aspect relationships, whereas most text summarization systems produce another unstructured text document.

B. SENTIMENT CLASSIFICATION AND SUBJECTIVE CLASSIFICATION

Sentiment analysis can be categorized into three subtasks: sentiment classification, subjective/objective identification, and aspect-based sentiment analysis. Sentiment classification, also known as document-level sentiment analysis, is the most broadly researched topic [6]–[10]. It classifies a review as conveying a positive or negative feeling. In this task, the whole document is considered as the elemental information unit. Researchers have shown that adjectives are good indicators of subjective and evaluative sentences [11]–[13]. Turney's group applied an unsupervised learning technique based on point-wise mutual information [14], [15], and Pang *et al.* used supervised machine learning methods (support vector machines (SVMs), naïve Bayes) to classify movie reviews [16]. Whitelaw *et al.* [7] applied WordNet to construct a lexicon. To automatically determine whether a term is indeed a marker of opinion content, Esuli and Sebastiani introduced SentiWordNet[1] as an enhanced lexical resource for sentiment analysis [17], and Ohana and Tierney applied SentiWordNet to document-level sentiment classification [18]. Recently, SentiWordNet 3.0 was released, with better accuracy than previous versions [19]. SentiStrength determines sentiment strength from informal English documents using a new method to exploit the de facto grammar and spelling styles of cyberspace [20]. Several researchers have used a lexicon-based approach to extract sentiments from text. A semantic orientation calculator performs the sentiment classification task using a dictionary of words tagged with their semantic orientations and incorporating intensification and negation [21]. Vo and Ock [22] proposed an unsupervised approach that classifies the polarity of a review with a combination of methods, including point-wise mutual information and SentiWordNet, and adjusts the phrase score in the case of modification. However, sentiment analysis at the document-level is too weak for most current applications.

The next level of sentiment analysis is subjective classification, which identifies subjective sentences. Sentence sentiment analysis usually includes two steps: identifying subjective sentences and classifying the opinions they express as positive or negative. Rilloff and Wiebe presented a bootstrapping approach that learns linguistic extraction patterns for sentiment expressions. A training set is automatically created by using a high-accuracy classifier to label a dataset, which is then taken as input by an extraction pattern-learning algorithm. The extracted patterns are then used to identify more and more subjective sentences. The pattern-learning algorithm learns many subjective patterns and progressively increases recall while retaining high precision [23]. Yu and Hatzivassiloglou [24] proposed several methods to identify sentence similarity, including the naïve Bayes classification. Mukund *et al.* used a modified SVM-based approach to distinguish subjective sentences from objective sentences [25].

Our work builds upon those developments in aspect-based sentiment analysis.

C. ASPECT-BASED SENTIMENT ANALYSIS

Opinion mining is valuable at both the document and sentence levels, but it does not determine precisely what people liked and disliked. Thus, algorithms are needed to digest a massive amount of information and extract product aspects and their corresponding opinions. In this research, we focus on identifying and extracting the product features that reviewers mention in their reviews. We considered several potential approaches to meet our goal.

1) EXTRACTION BASED ON FREQUENT NOUNS

Liu *et al.* used a data mining method to generate feature-based customer reviews [26], [27]. This algorithm detects explicit expressions (nouns and noun phrases) from a large review dataset. A part-of-speech tagger is applied to extract nouns and noun phrases. Their occurrence is calculated, and only frequently used ones are kept. This algorithm works because when reviewers comment on different features of a product, their words converge. The precision of this algorithm was improved in the Opine system [28], which uses relaxation labeling to identify the opinion orientation of words in context; this method accurately identifies sentiment phrases and their corresponding polarities. Moghaddam *et al.* improved the frequency-based method by adding a filter to delete non-aspect nouns [29]. Zhu *et al.* [31] introduced a technique that uses the C-value measure from [30] to identify multi-word aspects. Long *et al.* [32] proposed an aspect

[1] http://sentiwordnet.isti.cnr.it

extraction method based on frequency and information distance. Their system first identifies main feature words using the frequency-based method and then identifies other words related to the aspect using the information distance measure in [33].

2) EXTRACTION USING TOPIC MODELING

Topic modeling is an unsupervised machine learning approach used to summarize documents by considering each document as a mixture of topics and each topic as a probability distribution. Probabilistic latent semantic analysis (pLSA) [34] and latent Dirichlet allocation (LDA) are the two main techniques used for topic modeling [35]. In the opinion mining field, researchers have proposed a joint model to represent both sentiment words and topics simultaneously, which is possible because every opinion has a target. Mei *et al*. introduced an aspect sentiment mixture model using pLSA and a positive and negative sentiment training dataset. However, some researchers proved that topic modeling is unsuitable for identifying aspects (e.g., [36]). Later *et al*. introduced an approach that first identifies aspects using topic models and then detects opinion words by considering only adjectives [37]. In [10], a semi-supervised joint model enables the user to customize some seed feature terms for specific topics to generate aspect distributions that meet a specific requirement.

3) EXTRACTION USING OPINION TARGET RELATIONS

Opinions usually have a related target, so they can be mined to identify the aspects that are their targets using sentiment words. This approach was applied in [26] and [27] to extract aspects. For example, in the opinion "The software is amazing," "amazing" is an opinion word and "software" is the aspect. In [38]–[40], a DP was used to identify such relations for aspect extraction. In [38], Zhuang *et al*. proposed a dependency-based method for a movie review analysis application, and a double propagation method was proposed in [41] and [42] to iteratively exploit certain syntactic relationships between sentiment words and aspects.

4) EXTRACTION USING SUPERVISED LEARNING

Researchers have used many supervised learning approaches for subjective information extraction [43]–[45]. The most dominant approaches are based on sequential learning: hidden Markov models [46] and conditional random fields (CRFs) [47]. Yu *et al*. [48] introduced a supervised learning method called One-lass SVM [49] to identify aspects from the pros and cons of review format-2, as in [50]. The aspects are classified and ranked according to their frequency and their contributions to the overall rating of the reviews. Ghani *et al*. applied both semi-supervised learning and supervised learning for aspect identification [51]. Kovelamudi *et al*. introduced a supervised approach but also analyzed related data from Wikipedia [52]. Recently, Liu *et al*. introduced a supervised aspect identification method [53]. Their proposed system extracts aspects from various domains and retains the results as knowledge. The lifelong learning method clearly allows for better identification in a new domain than the original CRF. Recently, SemEval-2016 Task 5 released aspect-based sentiment analysis (ABSA) [54]. In [55], Toh and Su showed a two-component system: binary classifiers trained using a single layer feedforward network perform aspect category classification, and sequential labeling classifiers perform opinion target extraction. Meanwhile, Xenons *et al*.[56] introduced multiple ensembles based on SVMs for aspect category detection and a sequence labeling approach with CRFs for opinion target expression extraction.

Our task is related to but quite different from previous publications because we aim to build an automatic knowledge-based system by using NLP tools to extract coarse syntactic knowledge and infer opinion–aspect relationships from the statistics accumulated while obtaining the coarse knowledge. Our application will use opinion–aspect relationship knowledge, including product aspect inferences and sentiment extraction.

III. THE PROPOSED SYSTEM

Opinion mining at the document-level is the most widely used method for categorizing a whole-opinion review [14], [16]. Sentence-level sentiment analysis focuses on finding subjective sentences. In fact, most research has shown a close relationship between sentence- and document-level sentiment analyses [12], [14], [16], [57]. At both the document-level and the sentence-level, estimated opinion values are indirectly related to the topics (i.e., products or aspects of products) expressed in the text. They are useful, but they are too coarse for most applications because they do not identify the opinion targets. Without knowing what people liked and disliked, sentiment analysis is too narrow to be useful. In contrast, the aspect-based sentiment analyses found in recent surveys [10], [58]–[62] use more information from the review. To allow for an appropriate level of depth, we here emphasize a specific subfield of aspect-level analysis: product aspect extraction based on knowledge representation gained from reviews. To achieve fine-grained product aspect extraction, we detect and extract a customer's opinion about the individual parts of a product. Opinion values on individual aspects affect the aggregate opinion about a product to varying degrees. Our work thus addresses the issues of feature-based summaries of product reviews. In this paper, we focus on how to extract product aspects using the knowledge gained from reviews. However, before going into the details of the task, we need to define the terminology of our system.

Opinion: An opinion is defined as a quintuple $\left(e_j, a_{jk}, so_{ijkl}, h_i, t_l\right)$ [10], where e_j is a target entity (product), a_{jk} is an aspect of entity e_j (product aspect), so_{ijkl} is the sentiment score of the opinion held by opinion holder h_i about aspect a_{jk} of entity e_j at time t_l, h_i is an opinion holder, and t_l is the time when the opinion was expressed. For example, consider the camera review shown in Figure 1. The opinion holders "I" and "her" comment on the aspects "Canon G12 camera," I bought a Canon G12 camera six months ago. I simply love it. The picture quality is amazing. The battery life is also long. However, my wife thinks it is too heavy for her.

Opinion = (e_j, a_{jk}, so_{ijkl}, h_i, t_l)

Opinions =
{
Opinion 1 = {Canon G12 Camera, it/Canon G12 camera, simply_love, I, six months ago}
Opinion 2 = {Canon G12 Camera, picture quality, amazing, I, unknown}
Opinion 3 = {Canon G12 Camera, battery life, also_long, I, unknown}
Opinion 4 = {Canon G12 Camera, it/Canon G12, too_heavy, her, unknown}
}

FIGURE 1. An example application of the opinion model.

"picture quality," and "battery life" of the target entity "Canon G12 G3" using the sentiment values: "simply-love" for "Canon G12 camera," "amazing" for "picture quality," "also-long" for "battery life," and "however-too-heavy" on "Canon G12 Camera" at time "six months ago." We follow B. Liu's model [10] by focusing on opinion–aspect relation mining to extract product aspects and the related sentiments. The opinion holder and time are omitted because that information is not included in our dataset.

Aspect: An aspect is also known as a feature of the product that is the opinion target. The aspect can be one of these terms: a part of the given product, an attribute of the given product, or an attribute of a known aspect of the given product. A product can also be an aspect.

Opinion–Aspect Relationship: An opinion–aspect relationship denotes the relationship between a product aspect (opinion target) and a corresponding opinion.

Window: A window is the primary semantic unit, indicating terms and their relationships in a portion of a review. A window is made of a set of elements.

Element: An element is component of the window. An element is composed of an element key and an element value.

TABLE 1. Relations used in opinion–aspect relation extraction.

Relation	Description	Example
advmod	adverb modifier	it is too heavy
amod	adjective modifier	good price
ccomp	clausal complement	my wife thinks it is too heavy for her
cop	copula	it is too heavy
det	determiner	the picture quality is amazing
dobj	direct object	i simply love it
nn	noun compound modifier	the picture quality is amazing
npadvmod	noun phrase as adverbial modifier	six months ago
nsubject	nominal subject	the picture quality is amazing
num	numeric modifier	six months ago
pobj	object of a preposition	my wife thinks it is too heavy for her
poss	possession modifier	my wife thinks it is too heavy for her
prep	prepositional modifier	my wife thinks it is too heavy for her

Element Key: An element key is a binary relation identified from the parse tree by applying the DP tool. Table 1 lists the element keys used in our system. The whole set of relations is described in detail in the DP.[2]

[2]http://nlp.stanford.edu/software/stanford-dependencies.shtml

Element Value: An element value consists of a term from the text being analyzed and its annotations, as tagged by DP, CR, and NER.

Window Projection: A window projection is part of a window that can be found with regularity in numerous windows. Window projections are used to retain a particular interest for a specific purpose. For example, the section of a window that contains only an amod (adjective modifier) dependency relation is particularly useful for analyzing the opinion–aspect relation.

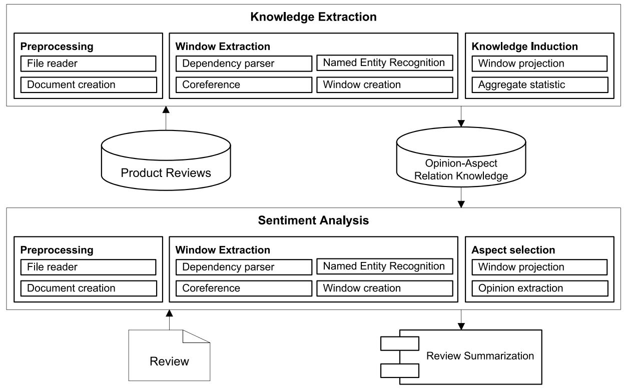

FIGURE 2. The proposed system.

The overall experimental procedure is illustrated in Figure 2. It involves two main stages: knowledge extraction and sentiment analysis. The knowledge extraction stage collects broad knowledge and implied opinion–aspect relationships from the reviews. In the second stage, that extracted knowledge is used to analyze new reviews and create an aspect-based summary. These stages are described in detail in the following sections.

A. KNOWLEDGE EXTRACTION FROM REVIEWS

Automatic knowledge extraction builds knowledge about opinion–aspect relationships. This section describes how they are captured from the raw text.

We focus on product reviews, which are highly focused on relevant information, such as opinions about a product. As discussed in the terminology, a product consists of a set of components and attributes called *aspects*. Moreover, the product itself is also an aspect. For example, a camera includes a set of components (e.g., screen, battery, lens), and a set of properties or attributes (e.g., image quality, options, weight). A screen also has its own set of attributes (e.g., screen quality, screen resolution, screen size). They are all product aspects about which opinions can be expressed. In this research, we paid close attention to the relationship between an opinion and its target. Normally, a reviewer provides opinions about specific targets, and those opinions can be expressed using a word, a part of a sentence, or a whole sentence in a review. To capture that sentiment information, we first formally define the model as a set of windows, i.e., $W = \{w_1, w_2, \ldots, w_n\}$, where n is the number of windows and $w_i = \{s_{i1}, s_{i2}, \ldots, s_{im}\}$ $(1 \leq i \leq n)$ is a window made of a set of m elements. We use $s_{ij} = \left\{k_{ij}, v_{ij}\right\}$ $(1 \leq j \leq m)$, where k_{ij} is an element key and v_{ij} is an element value, to denote element s_{ij} of window w_i.

In the first stage, our method takes a set of reviews as input and produces opinion–aspect relationship knowledge as output. We developed the system using a set of NLP tools: DP, CR,[3] and NER.[4] The creation process consists of three main phases.

1. **Preprocessing** — This is the preparation step.
2. **Window extraction** — Product reviews are taken as input and annotated by DP, CR, and NER, which perform relation detection and tagging. Windows are then filled with dependency relations and associated annotations. The outcome of window extraction is a set of raw windows representing broad knowledge obtained from the review documents.
3. **Window projection** — Window projection is applied to retain a particular interest (e.g., opinion–aspect relationships). Thus, projections of interest are determined from the windows, and a frequency for each projection is calculated. This step generates aggregate statistics from the obtained windows to infer opinion–aspect relationship knowledge.

The three steps of the opinion–aspect relationship knowledge creation process are illustrated in Figure 2 and presented in detail in the following sections.

1) PREPROCESSING

The system starts by reading a file that contains the reviews of customers adapted for use in the system. HTML tags and stop words are filtered. However, sentence structures are retained to extract the co-reference chains and opinion sequences. The aspect tags and sentiment labels are also stored in each document for evaluation.

2) WINDOW EXTRACTION

The important step in this process is the application of a set of NLP algorithms to identify the window elements for window extraction. First, a DP is applied to detect relationships between words within sentences to find opinion–aspect relation candidates. Second, CR is used to precisely identify the participating term or entity. Third, because opinion expressions in reviews are highly focused on a particular entity, NER is used to detect the types of terms on which a reviewer might comment. These processes increase the coverage of the window extraction when information is sparse. In addition, information gained by applying those NLP tools (DP, CR and NER) is then used in window elements to create intentional windows for specific purposes in the extensional parts of our system. For example, time, which is extracted by NER, is used to fill in the quintuple model of opinion. Syntactic feature annotations, such as degree, transitional, and negation words, can be used to adjust the sentiment values in the case of modifications and to detect the sequence of opinion in the whole document.

In this phase, the goal is to extract windows from the raw reviews. To extract the information of interest, window elements focus on expressed opinions. Each window includes a set of elements composed of element keys and element values. An element key is one of the binary relations listed in Table 1. A complex parse tree is isolated into several windows because each window can be only two levels deep. Depth limitation is required for the following reasons. First, given our focus on product aspect extraction, separating a parse tree forces each window to concentrate on a specific portion of text that expresses an opinion. Second, by restricting a window to be a subtree of a large parse tree, the system reduces the potential parser error in each window. Third, because the subtree isolating process also splits the sequence of opinions in the document, it prevents information omission by adding information annotated by the DP, CR, and NER to the windows. Table 2 illustrates how NER and CR annotations are added to windows extracted from the parse tree in Figure 3.

To fill the window elements, the input review is separated and dependency-parsed. Figure 3 illustrates the parse tree and graphical dependency representation for the example shown in Figure 1. The phrase structure and dependency representation are the two main varieties of syntactic annotation. In general, the phrase structure representation contains clear constituent structures and is suitable for language representation. In contrast, dependency representations can be better for languages largely independent of word order. By using the dependency representation, our model represents all relations uniformly as typed dependency relations in triplicate. For example, the dependency representation of the sentence, "The picture quality is amazing" is root (ROOT–0, amazing–5), det (quality–3, The–1), nn (quality–3, picture–2), nsubject (amazing–5, quality–3), cop (amazing–5, is–4), punct (amazing–5, .–6). The dependency is a binary relationship known as a grammatical relationship held between a head and a dependent. Opinion expression normally includes both an opinion and its target. Therefore, by applying a DP, we attempt to capture the relationship between opinions and product aspects such as nsubject (amazing–5, quality–3). Degree words (simply–love, too–heavy), transitional words (also–long), and negation words (however–thinks) are also extracted in this phase.

The DP helps to confirm the relationship, but it cannot capture a chain of opinions by linking terms in the whole document because of sentence and subtree isolation. For example, "it," which is used to avoid repetition in sentences, as in "I bought a Canon G12 camera six months ago. I simply love it" should be understood as the "Canon G12 camera." In this case, "Canon G12 camera" is known as the *antecedent* or *full form*, and "it" is an *anaphor* or *abbreviated form*. They should be interpreted as co–referential. In addition,

[3] https://stanfordnlp.github.io/CoreNLP/coref.html
[4] https://nlp.stanford.edu/software/CRF-NER.shtml

TABLE 2. Windows extracted from dependency parser in Figure 3.

Element key			Element value	NER and CR annotation	
Window 01					
$s_{[1,0]}$	root		bought/VBD		
$s_{[1,1]}$	nsubj		I/PRP		I
$s_{[1,2]}$	dobj		camera/NN		Canon G12 Camera
$s_{[1,3]}$		det	a/DT		Canon G12 Camera
$s_{[1,4]}$		nn	Canon/NNP	Org	Canon G12 Camera
$s_{[1,5]}$		nn	G12/NNP		Canon G12 Camera
$s_{[1,6]}$	advmod		Ago/RB	Time	
		npadvmod	Window 02		
Window 02					
$s_{[2,0]}$	root		months/NNS	Time	
$s_{[2,1]}$		num	Six/CD	Time	
Window 03					
$s_{[3,0]}$	root		love/VBP		I
$s_{[3,1]}$	nsubj		I/PRP		
$s_{[3,2]}$	dobj		it/PRP		Canon G12 Camera
$s_{[3,3]}$	advmod		simply/advmod		
Window 04					
$s_{[4,0]}$	root		amazing/JJ		
$s_{[4,1]}$	nsubj		quality/NN		
$s_{[4,2]}$		det	the/DT		
$s_{[4,3]}$		nn	picture/NN		
$s_{[4,4]}$	cop		is/VBZ		
Window 05					
$s_{[3,0]}$	root		long/JJ		
$s_{[5,1]}$	nsubj		life/NN		
$s_{[5,2]}$		det	the/DT		
$s_{[5,3]}$		nn	battery/NN		
$s_{[5,4]}$	cop		is/VBZ		
$s_{[5,5]}$	advmod		also/RB		
Window 06					
$s_{[6,0]}$	root		think/VBZ		
$s_{[6,1]}$	nsubj		wife/NN		my wife
$s_{[6,2]}$		poss	my/PRPS		I
$s_{[6,3]}$	ccomp		Window 07		
Window 07					
$s_{[7,0]}$	root		heavy		
$s_{[7,1]}$	nsubj		it/PRP		the battery life (Canon G12 Camera)
$s_{[7,2]}$	cop		is/VBZ		
$s_{[7,3]}$	advmod		too/RB		
$s_{[7,4]}$	prep		for/IN		
		pobj	her/PRP		my wife

TABLE 3. Co-reference information for the review in Figure 3.

Chain 1	"I" in sentence 1, i.e., 0–based character offsets [0, 1) "I" in sentence 2, i.e., 0–based character offsets [44, 45) "my" in sentence 5, i.e., 0–based character offsets [134, 136)
Chain 2	"a Canon G12 camera" in sentence 1, i.e., 0–based character offsets [9, 27) "it" in sentence 2, i.e., 0–based character offsets [58, 60)
Chain 3	"six months" in sentence 1, i.e., 0–based character offsets [28, 38)
Chain 4	"The picture quality" in sentence 3, i.e., 0–based character offsets [62, 81)
Chain 5	"The battery life" in sentence 4, i.e., 0–based character offsets [94, 110) "it" in sentence 5, i.e., 0–based character offsets [149, 151)
Chain 6	"my wife" in sentence 5, i.e., 0–based character offsets [134, 141) "her" in sentence 5, i.e., 0–based character offsets [169, 172)

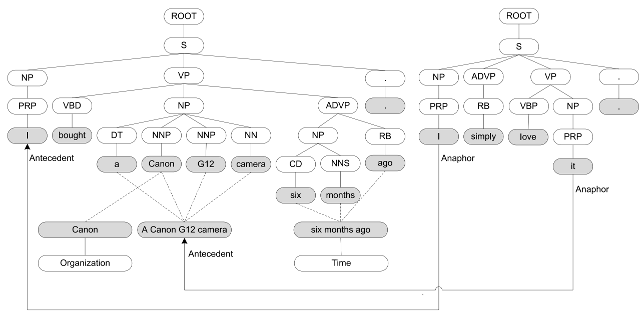

FIGURE 4. Graphical representation of CR and NER for a portion of the parse tree in Figure 3.

"Canon" is known as the name of a company, and "Canon G12 camera" is a product manufactured under the "Canon" brand. Therefore, we also use CR and NER in the system to obtain optimum performance. Table 3 and Figure 4 show how CR and NER are applied to a portion of the parser tree in Figure 3. Because none of the NLP systems are perfect, the CR sometimes delivers imprecise results. For example, Chain 5 shows that "it" (in sentence 5) is an abbreviated form of "The battery life." In fact, "it" should be understood as the "Canon G12 camera."

The output from window extraction is a massive set of windows signifying the broad knowledge obtained from the input review documents. These raw windows are taken as inputs for the knowledge induction step described next.

3) KNOWLEDGE INDUCTION

Knowledge induction comes from the large number of collected windows. The intentional knowledge patterns are obtained by exploiting redundancy in the reviews. For example, we intend to determine the product aspects and reviewers' opinions of them. Thus, the system first needs to isolate and examine particular portions of each window. In this specific case, we pay attention to the adjective → noun phrase relationships by which the system can infer opinion → aspect relationships. We expect to find the {adjective-noun} window very frequently, providing multiple windows that contain only that relation. We also derive reviewer actions from the {personal pronoun, verb} window. This kind of observation is used in the extensional parts of our system to analyze the relationship between opinions and actions.

To do this kind of induction analysis, we specify a projection operation on a window, which we call *window projection*. A window projection is a subset of windows that contain nonempty elements for a given subset of all elements. For example, we specify an A–N (adjective-noun) window projection. When the A-N projection is applied to a window, the system retains only the adjective (A) and noun (N) elements and ignores the rest. Similarly, we can specify window projections such as A-N-RB, where RB is an adverb that appends additional information related to the A–N window

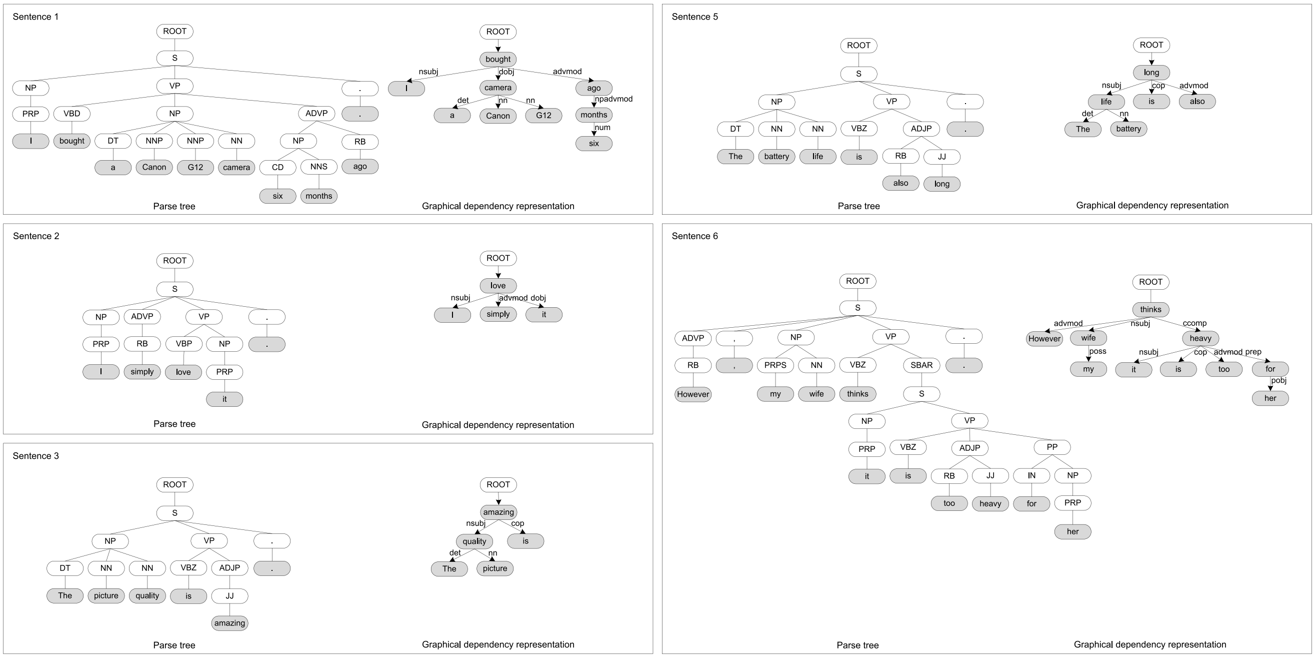

FIGURE 3. Parse tree and graphical dependency representation of the sentences "I bought a Canon G12 camera six months ago. I simply love it. The picture quality is amazing. The battery life is also long. However, my wife thinks it is too heavy for her."

TABLE 4. Example window projections from a portion of the parse tree in Figure 3.

A–N	amazing – quality/the picture quality	A $\xrightarrow[\text{nsubj}]{}$ N
A–N	long – life/the battery life	A $\xrightarrow[\text{nsubj}]{}$ N
A–N	it/a Canon G12 camera – heavy	A $\xrightarrow[\text{nsubj}]{}$ N
PRP–VBD–N	I – bought – camera/a Canon G12 camera	PRP $\xrightarrow[\text{nsubj}]{}$ VBD $\xleftarrow[\text{dobj}]{}$ N
PRP–VBD–N	I – love – it/a Canon G12 camera	PRP $\xrightarrow[\text{nsubj}]{}$ VBD $\xleftarrow[\text{dobj}]{}$ N

projection. Continuing with the example in Figure 3, we apply the A–N window projection to obtain a predominant noun that is a product aspect related to an adjective. The entailments rules were learned by using terminological axioms such as those in intentional windows. The PRP-VBD-PRP and PRP-VBD-NN patterns, which are shown in Table 4, represent the same projections due to the application of CR and NER. This kind of projection can also be applied to tasks such as opinion and action mining. Window projections analyze windows along multiple dimensions and allow us to efficiently precompute and look up aggregate statistics.

Frequency and conditional probability are two valuable aggregate statistics that can be obtained from the window projection. The frequency of a window (or window element) can be computed as the number of other windows (elements) whose elements match with all the elements of a given window. The frequency statistic provides an evaluation of the popularity of patterns in a given set of elements. For example, a high frequency of {adjective–noun} windows could indicate that the noun is an aspect of the product being reviewed. However, the frequency value depends on the size of the dataset, as discussed below. The second aggregate statistic is the conditional probability of a specific window; it can be calculated from the probability of a specific set of patterns given a subset of element patterns or particular information. For example, the probability for {adjective, noun} windows in which the adjective = small is {0.8 for {small, camera}, 0.2 for {small–card}}, which indicates that the nouns modified by the adjective *small* are camera and card with the probability of 0.8 and 0.2, respectively. Similarly, by calculating the probabilities for {adjective, noun} windows in which the noun = quality produces the result that the adjective before the noun will be high or light, with probabilities of 0.9 and 0.1, respectively.

B. SENTIMENT ANALYSIS

At this stage, we present a potential application that can use opinion–aspect relationship knowledge, such as in a product aspect inference or sentiment extraction. The system takes a new product review as input and treats the review in the same way as in the first stage. However, instead of performing knowledge induction, the system performs opinion extraction. First, it dissects and performs a multidimensional analysis by applying the preprocessing and window extraction process steps. Second, it uses the opinion–aspect relation knowledge to extract opinions about product aspects using opinion words and modifiers such as degree, transitional, and negation words. Third, it combines those data to produce an aspect-based summary of the review.

1) PRODUCT ASPECT INFERENCE

This is most important task in our system. It extracts the product aspects to which the opinions refer. As shown in the section on window projections, the system dissects windows along different dimensions. We then use the aggregate statistics from across all the training reviews, our aspect–opinion relationship knowledge, to determine candidate product aspects. Given a new review, the system can use already extracted knowledge to infer product aspects. For example, the aggregate statistics suggest that the most common noun phase in relation with an adjective is the product aspect. Therefore, the system can infer that in the sentence "The picture quality is amazing," "picture quality" is the most likely aspect candidate. Similarly, from the sentence "The battery life is also long," the system learns from a large amount of data that "long" is typically expressed in regard to the battery aspect and thus correctly infers "battery life" as a product aspect.

2) SENTIMENT DETECTION

As shown above, inferences about product aspects also indicate sentiment words expressed about those aspects. For example, in the sentence, "The picture quality is amazing," "amazing" can be recognized as the opinion and "picture quality" as the target. Thus, the system simultaneously extracts both "amazing" and "picture quality," then fills in the quintuple model of opinion. After that, SentiWordNet[5] is applied to automatically verify whether the extracted phrase is indeed a marker of sentiment content. Each synset in SentiWordNet receives three numerical values, Obj (objective), Pos (positive), and Neg (negative), that indicate the sentiment score for the terms it contains. Thus, our system determines the sentiment value of each extracted phrase using a sentiment score derived from SentiWordNet.

3) SYNTACTIC FEATURE EXTRACTION

When reviewers express an opinion about an aspect, they frequently use modifier words to alter the meaning of the opinion. In most cases, the modifier provides additional information about another word in the sentence. Consequently, we consider not only opinion words, but also their modifiers. Modifiers can include degree, negation, and transitional words. Degree words (e.g., very, quite, and too) add meaning to a word or phrase. For example, the expression "It is too heavy for her," contains the modifier "too," which is denoted by the advmod (too, heavy) relation. If "heavy" is moderately negative, and then "too heavy" indicates a strongly negative expression. Similarly, negation words (e.g., not, no, and never) invert the meaning of a word, phrase, or clause in a sentence. Figure 5 shows an example of how the system deals with the influence of modifiers by adjusting the opinion value.

[5]http://sentiwordnet.isti.cnr.it

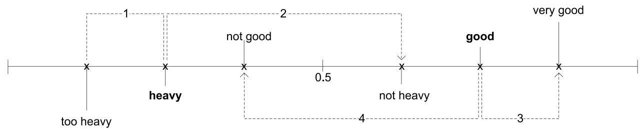

FIGURE 5. Example of adjusting the opinion in cases of modification. In the camera domain, assume that *heavy* is strongly negative and *good* is moderately positive. (1) too heavy: very strongly negative (2) not heavy: moderately positive (3) very good: very strongly positive (4) not good: moderately negative.

4) IMPLICIT OPINION INFERENCE AND SEQUENCE OF OPINION

As discussed in the previous section, a review typically contains multiple opinions, which are called a *sequence of opinions*. Consider again the product review: "I bought a Canon G12 camera six months ago. I simply love it. The picture quality is amazing. The battery life is also long. However, my wife thinks it is too heavy for her." The first sentence contains only a fact and offers no opinion. The remaining sentences all contain either implicit or explicit opinions. In some cases, the opinions are linked by transitional words (e.g., also, and, however), which are used to find sentiment in an implicit expression or correctly infer sentiment in a sequence of opinions. Figure 6 shows how the transitional words change the opinion polarity. Assume that "picture quality is amazing" is positive. The word "also" normally implies same-idea transitions; thus we can correctly infer that the "battery life is long" is positive. Similarly, the word "however", which begins the last sentence, suggests that the opinion that the camera is heavy has negative polarity.

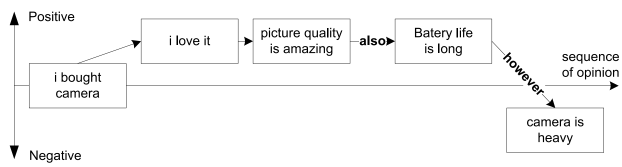

FIGURE 6. Example of adjusting the opinion in cases of modification.

5) REVIEW SUMMARIZATION

As discussed in the terminology section, the opinion quintuple is designed to give some general information for an opinion summary. In this work, the opinion holder and time are omitted because that information is not included in our dataset. In our case, the summary is generated based on product aspects, so we call it the *aspect-based sentiment analysis*. Figure 7 illustrates this summary for the above single review.

Canon G12 Camera

Aspect	Positive	Negative
Camera	1 <i simply love canon_G12_camera>	
Picture Quality	1 <picture quality is amazing>	
Battery Life	1 <batery life is also long>	
Weight		1 <canon_G12_camera/weight is too heavy>

FIGURE 7. Example of a single review summary.

The summarization process can be applied to an individual review or a set of product reviews. Thus, we can achieve a summary of opinions in a structured manner by analyzing the opinions of a large number of customers.

IV. EVALUATION AND DISCUSSION

A. EVALUATION

In general, component-level and system-level metrics are commonly used to evaluate a knowledge-base system. Typically, a component-level evaluation measures the quality of the knowledge base independent of particular applications. A system-level evaluation can be derived from the component level information about an individual application. To evaluate the performance of our aspect extraction, we use a component-level metric to evaluate the correctness and coverage of the system.

TABLE 5. Datasets.

	Camera		Laptop	
Measure	Training (Canon G3)	Test (Nikon 4300)	Training set	Test set
No. of reviews	45	34	64	16
No. of sentences	597	364	614	194
No. of sentences per review	13	11	10	12
No. of windows extracted	1764	948	1353	497
No. of windows per sentence	2.95	2.60	2.20	2.56
Maximum depth of graphical dependency representation	10	9	10	8
No. of relations extracted	9483	5398	7484	2759
No. of desired aspect–opinion relations	316	203	621	180

1) COVERAGE

A large number of sources is needed for a knowledge-based system. We used customer review data[6] and the SemEval-2016 Laptop Reviews-English[7] to test our new system and evaluate its performance. The customer review dataset contains annotated customer reviews of five products from amazon.com. They are labeled with respect to product aspect and opinion. The SemEval-2016 Laptop dataset was distributed in the context of SemEval-2016 and has annotated aspect categories and sentiment polarity labels at the text level. Table 5 summarizes the details of our experimental data. There were 961 sentences in the Camera domain and 808 sentences in the Laptop domain. The dataset contains fairly long reviews, with 11.44 sentences per review on average. Our system's coverage evaluation results are shown in Table 5 and Table 6. The coverage experiment gave us an estimate of the quality of the system; however the system's effects on a specific application are more important.

[6]https://www.cs.uic.edu/∼liub/FBS/sentiment-analysis.html#datasets
[7]http://alt.qcri.org/semeval2016/task5/index.php?id=data-and-tools

TABLE 6. Relations extracted in window extraction step.

Type	Description	Number of relations extracted			
		Training	Test	Training	Test
Total	Total relations extracted	9483	5398	7484	2759
amod	No. of adjective modifier relations	771 (8.13%)	442 (8.18%)	416 (5.56%)	145 (5.26%)
nsubject	No. of nominal subject relations	1148 (12.11%)	691 (12.80%)	994 (13.28%)	374 (13.56%)
dobj	No. of direct object relations	657 (6.93%)	385 (7.13%)	520 (6.95%)	199 (7.21%)
advmod	No. of adverb modifier relations	573 (6.04%)	329 (6.09%)	517 (6.91%)	180 (6.52%)
Others		6334 (66.79%)	3551 (65.78%)	5037 (67.30%)	1861 (67.45)

Table 6 shows that the system used approximately 33% of the extracted relationships. Thus, many potential relationships could improve the performance of the system. Table 7 lists the most common aspects, opinions, and verbs. The Camera and Laptop domains share common aspects; however, each domain contains various aspects specific to each product. The extracted verbs are used to link sentiments and used in behavior models in extensional parts of our system.

TABLE 7. Common aspects, opinions, and verbs.

Camera			Laptop		
aspects	opinions	verbs	aspects	opinions	verbs
camera	great	have	laptop	great	buy
flash	new	buy	computer	new	have
picture	easy	want	keyboard	easy	want
image	nice	use	disk	nice	use
photo	fast	get	system	fast	get
quality	good	purchase	display	high	purchase
use	high	go	port	best	find
card	extra	find	price	impressed	buy
mode	hard	work	battery	heavy	work
software	best	take	product	simple	need
lens	light	recommend	machine	amazing	slow
setting	impressed	see	store	small	take
feature	heavy	pay	mouse	fantastic	recommend
choice	simple	ask	setting	large	run
option	old	consider	value	long	consider
control	slow	return	network	awesome	return
battery	amazing	think	feature	bad	think
button	small	decide	quality	expensive	decide
screen	fantastic	make	size	excellent	expect
problem	big	expected	wireless	happy	take

2) CORRECTNESS

In the camera domain, we used the Canon G3 dataset for training and the Nikon 4300 reviews for testing. In the laptop domain, we used 80% of the reviews in the SemEval-2016 Laptop dataset for training and the remaining 20% for testing. We used precision, recall, and F1-score as evaluation measures. As shown in Table 8, we achieved satisfactory

TABLE 8. Aspect extraction results for precision, recall, and f1 scores: camera and laptop reviews.

	Camera	Laptop
No. of actual aspects	203	180
No. of aspects extracted	262	205
True positives	166	149
False positives	96	56
False negatives	37	31
Precision	0.634	0.773
Recall	0.818	0.828

experimental results, with especially good results in the Laptop domain.

Aspect extraction was evaluated using camera reviews, and it showed an F1-score of 0.714. The evaluation of aspect extraction based on laptop reviews showed an F1-score of 0.774. A comparison between our results and those from related works is given in Table 9.

TABLE 9. Comparison with related works on camera reviews.

	Camera		
	P	R	F_1
Our experiment	0.634	0.818	0.714
Frequent features (Hu and Liu, 2004)	0.552	0.671	0.606
Infrequent feature identification (Hu and Liu, 2004)	0.747	0.822	0.783
Cross-domain lifelong learning CRF(Hu and Liu, 2017)	0.819	0.606	0.696
In-domain lifelong learning CRF(Hu and Liu, 2017)	0.807	0.754	0.779

B. DISCUSSION AND ANALYSIS

1) RESULT DISCUSSION

Liu *et al*. analyzed the same dataset for product aspect extraction using several methods and obtained the best performance using the lifelong learning CRF [53]. Table 9 shows the comparison between our experimental results and related works in the camera domain. Liu *et al*.'s lifelong learning method clearly allows for better identification in a new domain than the original CRF. Our experiment also showed that the Camera and Laptop domains share common aspects, such as a screen, battery, and settings. However, we focused on a specific domain, obtaining knowledge by performing aspect extraction on many reviews in an individual domain. Although internet reviews are dispersed in many forms and domains, the acquisition of specific knowledge from a domain or product series is valuable and profitable. For example, by collecting review datasets on the Samsung Galaxy S series (i.e., S5, S6, S7), we could automatically obtain the knowledge in those datasets and apply it in future problem solving for S8 review mining. This work would be profitable because of the many aspects shared across the product series. Table 10 shows the results of related work that shared subtasks for aspect-based sentiment analysis [54]. Using the same laptop dataset, those researchers performed the task of identification on the entity and attributes (subtask 1, slot 1). In their case, the entity and attributes were chosen from a predefined list, such as laptop, display, and general. They achieved satisfactory results, and the best performance was obtained using a binary classifier, which was trained using a single-layer feedforward network. Our system automatically obtains the product aspects from the reviews. We perform the broad task of raw aspect extraction; thus, we relabeled the dataset to fit our evaluation. Furthermore, the structured knowledge representation gained from reviews can be used for deeper inferences. We can exploit such group aspects by mining various words and phrases that aim to represent a single product aspect. For example, in our experiment we found that screen, display, and LCD all refer to the same aspect of the camera. Similarity, disk, HDD, and SSD refer to the same feature of a laptop. Thus, our system can leverage a large dataset to initialize a deeper knowledge from aspect extraction.

TABLE 10. Results of related works on laptop reviews.

	Laptop		
	P	R	F_1
Our experiment	0.773	0.828	0.774
Neural network features, slot 1 (NLANGP, 2016)	0.569	0.478	0.519
Ensembles of classifiers and embeddings, slot 1 (AUEB, 2016)	0.456	0.532	0.491

2) FACING PROBLEMS

Because reviews contain subjective and objective information, we must consider parser error, writing style, incomplete sentences, short text, grammatical errors, informal text, internet slang, emotions, comparisons, domain dependence, sarcastic statements, and other informalities. We improved the system to address those issues. First, by isolating each subtree, we focus on an immediate opinion expression about an aspect. Furthermore, by restricting each window to contain only a small subtree from a large parse tree, we can change a parser error in one window without removing information. The second area of improvement is adding information annotated by DP, CR, and NER. This provides extra information, including context, reference, and domain, that we use to improve aspect extraction and sentiment classification. Finally, we might be able to produce a better result by incorporating larger datasets. Although we applied a suite of NLP components in our system, aspect extraction remains a difficult task. We intend to gain inferential power by linking opinion and behavior models. In that new model, an opinion could be based on facts, but it is subjective and could also be based on feelings or judgments. Both facts and opinions are important parts of reviews. We aim to determine the effects of behavior on sentiment, extract the indicators of opinion, and analyze the sequence of opinion in the whole document. Because no NLP system is perfect, we aim to improve the accuracy of our NLP system by initializing an adequately deep knowledge of the product aspects.

V. CONCLUSION

This paper addresses product-aspect-extraction-based knowledge in product reviews. We introduce a system that works in two main stages: knowledge extraction and sentiment analysis. First, the system automatically extracts broad syntactic knowledge and infers opinion–aspect relationships using the DP, CR, and NER NLP tools. The knowledge creation process isolates subtrees, extracts dependency relations, and detects additional annotations, such as co-reference chains, named entity annotations, and syntactic features. Second, that knowledge is used to analyze new reviews and generate a feature-based summary. Product aspect extraction was performed and achieved satisfactory experimental results, especially for the Laptop domain. This strategy could offer a new approach to product aspect extraction in large datasets and potentially be applied to challenging tasks such as implicit opinion inferences, sarcastic statements, and the opinion-behavior model.

REFERENCES

[1] G. Salton, A. Singhal, C. Buckley, and M. Mitra, "Automatic text decomposition using text segments and text themes," in *Proc. HYPERTEXT*, Bethesda, MD, USA, 1996, pp. 53–65.

[2] J. Goldstein, M. Kantrowitz, V. Mittal, and J. Carbonell, "Summarizing text documents: Sentence selection and evaluation metrics," in *Proc. SIGIR*, Berkeley, CA, USA, 1999, pp. 121–128.

[3] C. D. Paice, "Constructing literature abstracts by computer: Techniques and prospects," *Inf. Process. Manage.*, vol. 26, no. 1, pp. 171–186, 1990.

[4] J. Kupiec, J. O. Pedersen, and F. Chen, "A trainable document summarizer," in *Proc. SIGIR*, Seattle, WA, USA, 1995, pp. 68–73.

[5] O. Gross, A. Doucet, and H. Toivonen, "Document summarization based on word associations," in *Proc. SIGIR*, Gold Coast, QLD, Australia, 2014, pp. 1023–1026.

[6] K. Dave, S. Lawrence, and D. M. Pennock, "Mining the peanut gallery: Opinion extraction and semantic classification of product reviews," in *Proc. WWW*, Budapest, Hungary, 2003, pp. 519–528

[7] C. Whitelaw, N. Garg, and S. Argamon, "Using appraisal groups for sentiment analysis," in *Proc. CIKM*, Bremen, Germany, 2005, pp. 625–631.

[8] S. R. Das and M. Y. Chen, "Yahoo! For Amazon: Sentiment extraction from small talk on the Web," *Manage. Sci.*, vol. 53, no. 9, pp. 1375–1388, 2007.

[9] A. Devitt and K. Ahmad, "Sentiment polarity identification in financial news: A cohesion-based approach," in *Proc. ACL*, Prague, Czech Republic, 2007, pp. 984–991.

[10] B. Liu, "Sentiment analysis and opinion mining," *Synth. Lectures Human Lang. Technol.*, vol. 5, no. 1, pp. 1–167, 2012.

[11] V. Hatzivassiloglou and J. M. Wiebe, "Effects of adjective orientation and gradability on sentence subjectivity," in *Proc. COLING*, Saarbrücken, Germany, 2000, pp. 299–305.

[12] J. M. Wiebe, "Learning subjective adjectives from corpora," in *Proc. AAAI/IAAI*, Menlo Park, CA, USA, 2000, pp. 735–740.

[13] J. Wiebe, T. Wilson, and M. Bell, "Identifying collocations for recognizing opinions," in *Proc. ACL/EACL*, Toulouse, France, 2001, pp. 24–31.

[14] P. D. Turney, "Thumbs up or thumbs down?: Semantic orientation applied to unsupervised classification of reviews," in *Proc. ACL*, Philadelphia, PA, USA, 2002, pp. 417–424.

[15] P. D. Turney and M. L. Littman. (2002). "Unsupervised learning of semantic orientation from a hundred-billion-word corpus." [Online]. Available: https://arxiv.org/abs/cs/0212012

[16] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up?: Sentiment classification using machine learning techniques," in *Proc. EMNLP*, Philadelphia, PA, USA, 2002, pp. 79–86.

[17] A. Esuli and F. Sebastiani, "SENTIWORDNET: A publicly available lexical resource for opinion mining," in *Proc. LREC*, Genoa, Italy, 2006, pp. 417–422.

[18] B. Ohana and B. Tierney, "Sentiment classification of reviews using SentiWordNet," in *Proc. IT&T Conf.*, Dublin, Ireland, 2009.

[19] S. Baccianella, A. Esuli, and F. Sebastiani, "SentiWordNet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining," in *Proc. LREC*, vol. 10. 2010, pp. 2200–2204.

[20] M. Thelwall, K. Buckley, G. Paltoglou, D. Cai, and A. Kappas, "Sentiment strength detection in short informal text," *J. Assoc. Inf. Sci. Technol.*, vol. 61, no. 12, pp. 2544–2558, 2010.

[21] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede, "Lexicon-based methods for sentiment analysis," *Comput. Linguistics*, vol. 37, no. 2, pp. 267–307, 2011.

[22] A.-D. Vo and C.-Y. Ock, "Sentiment classification: A combination of PMI, SentiWordNet and fuzzy function," in *Proc. ICCCI*, Ho Chi Minh City, Vietnam, 2012, pp. 373–382.

[23] E. Riloff and J. Wiebe, "Learning extraction patterns for subjective expressions," in *Proc. EMNLP*, 2003, pp. 105–112.

[24] H. Yu and V. Hatzivassiloglou, "Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences," in *Proc. EMNLP*, 2003, pp. 129–136.

[25] S. Mukund and R. K. Srihari, "A vector space model for subjectivity classification in Urdu aided by co-training," in *Proc. COLING*, Beijing, China, 2010, pp. 860–868.

[26] M. Hu and B. Liu, "Mining and summarizing customer reviews," in *Proc. KDD*, Seattle, WA, USA, 2004, pp. 168–177.

[27] M. Hu and B. Liu, "Mining opinion features in customer reviews," in *Proc. AAAI*, San Jose, CA, USA, 2004, pp. 755–760.

[28] A.-M. Popescu and O. Etzioni, "Extracting product features and opinions from reviews," in *Proc. HLT EMNLP*, Vancouver, BC, Canada, 2005, pp. 339–346.

[29] S. Moghaddam and M. Ester, "ILDA: Interdependent LDA model for learning latent aspects and their ratings from online product reviews," in *Proc. SIGIR*, Beijing, China, 2011, pp. 665–674.

[30] K. T. Frantzi, S. Ananiadou, and J. I. Tsujii, "The *C-value/NC-value* method of automatic recognition for multi-word terms," in *Proc. ECDL*, Heraklion, Greece, 1998, pp. 585–604.

[31] J. Zhu, H. Wang, B. K. Tsou, and M. Zhu, "Multi-aspect opinion polling from textual reviews," in *Proc. CIKM*, Hong Kong, 2009, pp. 1799–1802.

[32] C. Long, J. Zhang, and X. Zhut, "A review selection approach for accurate feature rating estimation," in *Proc. COLING*, Beijing, China, 2010, pp. 766–774.

[33] R. L. Cilibrasi and P. M. B. Vitanyi, "The Google similarity distance," *IEEE Trans. Knowl. Data Eng.*, vol. 19, no. 3, pp. 370–383, Mar. 2007.

[34] T. Hofmann, "Probabilistic latent semantic indexing," in *Proc. SIGIR*, Berkeley, CA, USA, 1999, pp. 50–57.

[35] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," *J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, Mar. 2003.

[36] I. Titov and R. McDonald, "Modeling online reviews with multi-grain topic models," in *Proc. WWW*, Beijing, China, 2008, pp. 111–120

[37] S. Brody and N. Elhadad, "An unsupervised aspect-sentiment model for online reviews," in *Proc. NAACL HLT*, Los Angeles, CA, USA, 2010, pp. 804–812.

[38] L. Zhuang, F. Jing, and X.-Y. Zhu, "Movie review mining and summarization," in *Proc. CIKM*, Arlington, VA, USA, 2006, pp. 43–50.

[39] S. Somasundaran, J. Ruppenhofer, and J. Wiebe, "Discourse level opinion relations: An annotation study," in *Proc. SIGdial*, Columbus, OH, USA, 2008, pp. 129–137.

[40] N. Kobayashi, R. Iida, K. Inui, and Y. Matsumoto, "Opinion mining on the Web by extracting subject-aspect-evaluation relations," in *Proc. AAAI-CAAW*, 2006, pp. 86–91.

[41] G. Qiu, B. Liu, J. Bu, and C. Chen, "Expanding domain sentiment lexicon through double propagation," in *Proc. IJCAI*, Pasadena, CA, USA, 2009, pp. 1199–1204.

[42] L. Zhang, B. Liu, S. H. Lim, and E. O'Brien-Strain, "Extracting and ranking product features in opinion documents," in *Proc. COLING*, Beijing, China, 2010, pp. 1462–1470.

[43] *Handbook of Natural Language Processing*, Marcel Dekker Inc., New York, NY, USA, 2000.

[44] J. L. Leidner, "Book review: Handbook of natural language processing," *Comput. Linguistics*, vol. 37, no. 2, pp. 395–397, 2011.

[45] R. J. Mooney and R. Bunescu, "Mining knowledge from text using information extraction," *ACM SIGKDD Explorations Newslett.*, vol. 7, no. 1, pp. 3–10, 2005.

[46] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, no. 2, pp. 267–296, Feb. 1990.

[47] J. D. Lafferty, A. McCallum, and F. C. N. Pereira, "Conditional random fields: Probabilistic models for segmenting and labeling sequence data," in *Proc. ICML*, vol. 1. 2001, pp. 282–289.

[48] J. Yu, Z.-J. Zha, M. Wang, and T.-S. Chua, "Aspect ranking: Identifying important product aspects from online consumer reviews," in *Proc. ACL-HLT*, Portland, OR, USA, 2011, pp. 1496–1505.

[49] L. M. Manevitz and M. Yousef, "One-class SVMs for document classification," *J. Mach. Learn. Res.*, vol. 2, pp. 139–154, Dec. 2001.

[50] B. Liu, M. Hu, and J. Cheng, "Opinion observer: Analyzing and comparing opinions on the Web," in *Proc. WWW*, Chiba, Japan, 2005, pp. 342–351

[51] R. Ghani, K. Probst, Y. Liu, M. Krema, and A. Fano, "Text mining for product attribute extraction," *SIGKDD Explorations Newslett.*, vol. 8, no. 1, pp. 41–48, 2006.

[52] S. Kovelamudi, S. Ramalingam, A. Sood, and V. Varma, "Domain independent model for product attribute extraction from user reviews using Wikipedia," in *Proc. IJCNLP*, Chiang Mai, Thailand, 2011, pp. 1408–1412.

[53] L. Shu, H. Xu, and B. Liu. (2017). "Lifelong learning CRF for supervised aspect extraction." [Online]. Available: https://arxiv.org/abs/1705.00251

[54] M. Pontiki *et al.*, "SemEval-2016 task 5: Aspect based sentiment analysis," in *Proc. SemEval*, San Diego, CA, USA, 2016, pp. 19–30.

[55] Z. Toh and J. Su, "NLANGP at SemEval-2016 task 5: Improving aspect based sentiment analysis using neural network features," in *Proc. SemEval*, San Diego, CA, USA, 2016, pp. 282–288.

[56] D. Xenos, P. Theodorakakos, J. Pavlopoulos, P. Malakasiotis, and I. Androutsopoulos, "AUEB-ABSA at SemEval-2016 task 5: Ensembles of classifiers and embeddings for aspect based sentiment analysis," in *Proc. SemEval*, San Diego, CA, USA, 2016, pp. 312–317.

[57] B. Pang and L. Lee, "A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts," in *Proc. ACL*, Barcelona, Spain, 2004, p. 271.

[58] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Found. Trends Inf. Retr.*, vol. 2, nos. 1–2, pp. 1–135, 2008.

[59] M. Tsytsarau and T. Palpanas, "Survey on mining subjective data on the Web," *Data Mining Knowl. Discovery*, vol. 24, no. 3, pp. 478–514, May 2012.

[60] H. Tang, S. Tan, and X. Cheng, "A survey on sentiment detection of reviews," *Expert Syst. Appl.*, vol. 36, no. 7, pp. 10760–10773, Sep. 2009.

[61] K. Ravi and V. Ravi, "A survey on opinion mining and sentiment analysis: Tasks, approaches and applications," *Know. Based Syst.*, vol. 89, pp. 14–46, Nov. 2015.

[62] W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications: A survey," *Ain Shams Eng. J.*, vol. 5, no. 4, pp. 1093–1113, Dec. 2014.

ANH-DUNG VO received the B.S. degree in computer science from the Hanoi University of Science and Technology, Vietnam, in 2010, and the M.S. degree in information technology from the University of Ulsan, South Korea, in 2013. He is currently pursuing the Ph.D. degree with the University of Ulsan.

His research interests include natural language processing, machine learning, and sentiment analysis.

QUANG-PHUOC NGUYEN received the B.S. degree in information technology from Vietnam National University–HCM City, University of Natural Sciences, Vietnam, in 2005, and the M.S. degree in information technology from Konkuk University, South Korea, in 2010. He is currently pursuing the Ph.D. degree with the University of Ulsan, South Korea.

His research interests include natural language processing, machine learning, and machine translation.

CHEOL-YOUNG OCK received the B.S., M.Sc., and Ph.D. degrees in information technology from Seoul National University, South Korea, in 1982, 1984, and 1993, respectively. He was the Visiting Professor at Tomsk Polytechnic University, Russia, in 1994, and the University of Glasgow, U.K., in 1996. He is currently a Professor with the University of Ulsan, South Korea.

His research interests include natural language processing, ontology, information retrieval, and machine learning.