Feature and Opinion Mining for Customer Review Summarization

Muhammad Abulaish[1,*], Jahiruddin[1], Mohammad Najmud Doja[2], and Tanvir Ahmad[2]

[1] Department of Computer Science, Jamia Millia Islamia, New Delhi, India
abulaish@ieee.org, jahir.jmi@gmail.com
[2] Department of Computer Engineering, Jamia Millia Islamia, New Delhi, India
ndoja@yahoo.com, tanvir.ce@jmi.ac.in

Abstract. In this paper, we present an opinion mining system to identify product features and opinions from review documents. The features and opinions are extracted using semantic and linguistic analysis of text documents. The polarity of opinion sentences is established using polarity scores of the opinion words through Senti-WordNet to generate a feature-based summary of review documents. The system is also integrated with a visualization module to present feature-based summary of review documents in a comprehendible way.

Keywords: Opinion mining, Opinion analysis, Sentiment analysis, Text mining, Review summarization, Natural language processing.

1 Introduction

In recent past, due to existence of numerous forums, discussion groups, and blogs, individual users are participating more actively and are generating vast amount of new data – termed as *user-generated contents*. These new Web contents include customer reviews and blogs that express opinions on products and services – which are collectively referred to as customer feedback data on the Web. As customer feedback on the Web influences other customer's decisions, these feedbacks have become an important source of information for businesses to take into account when developing marketing and product development plans.

Recent works have shown that the distribution of an overwhelming majority of reviews posted in online markets is bimodal. Reviews are either allotted an extremely high rating or an extremely low rating. In such situations, the average numerical star rating assigned to a product may not convey a lot of information to a prospective buyer. Instead, the reader has to read the actual reviews to examine which of the positive and which of the negative aspect of the product are of interest. Several sentiment analysis approaches have proposed to tackle this challenge up to some extent. However, most of the classical sentiment analysis mapping the customer reviews into binary classes – *positive* or *negative*, fails to identify the product features liked or disliked by the customers.

* To whom correspondence should be addressed.

S. Chaudhury et al. (Eds.): PReMI 2009, LNCS 5909, pp. 219–224, 2009.

In this paper, we present an opinion mining system which uses linguistic and semantic analysis of text to identify key information components from text documents. The information components are centered on both product features, and associated opinions, which are extracted using natural language processing techniques and co-occurrence-based analysis. The novelty of the system lies in mining associated modifiers with opinions to represent the degree of expressiveness of opinions. For each extracted feature, the list of opinions and associated modifiers are compiled and their polarity is established using numerical scores obtained through Senti-WordNet [8]. We also present a visualization technique that provides a feature-based summary of review documents in a graphical way. The feature-based summary can help the customers as well as manufacturers to know about the positive and negative aspects of the products without going through pile of documents.

The remaining paper is structured as follows: Section 2 presents related works on opinion mining. Section 3 presents the architectural details of proposed opinion mining system. The evaluation of the feature and opinion extraction process is presented in section 4. Finally, section 5 concludes the paper with possible enhancements to the proposed system.

2 Related Work

Research on opinion mining started with identifying opinion bearing words, e.g., *great*, *amazing*, *wonderful*, *bad*, *poor* etc. Many researchers have worked on mining such words and identifying their semantic orientations. In [3], a bootstrapping approach is proposed, which uses a small set of given seed opinion words to find their synonyms and antonyms in WordNet. The history of the phrase *sentiment analysis* parallels that of *opinion mining* in certain respects. A sizeable number of papers mentioning *sentiment analysis* focus on the specific application of classifying customer reviews as to their polarity – *positive* or *negative* [4,6]. Although, classical sentiment classification attempts to assign the review documents either positive or negative class, it fails to find what the reviewer or opinion holder likes or dislikes. To obtain detailed aspects, feature-based opinion mining is proposed in literature [1,3,5]. In [1], a supervised pattern mining method is proposed. In [3,5], an unsupervised method is used. A lexicon-based approach has been shown to perform quite well in [2,3]. The lexicon-based approach basically uses opinion words and phrases in a sentence to determine the orientation of an opinion on a feature.

Although, some opinion mining methods extract features and opinions from document corpora, most of them do not explicitly exploit the semantic relationships between them. The proposed method differs from all these approaches predominantly in its use of pure linguistic techniques to identify only those features for which customers have commented using opinionated words. Moreover, extraction of associated modifiers used in review documents to represent the degree of expressiveness of opinions is unique in our work.

3 Proposed Opinion Mining System

Fig. 1 presents the architectural details of the proposed opinion mining system, which consists of five major modules – *Document Processor*, *Subjectivity/ Objectivity*

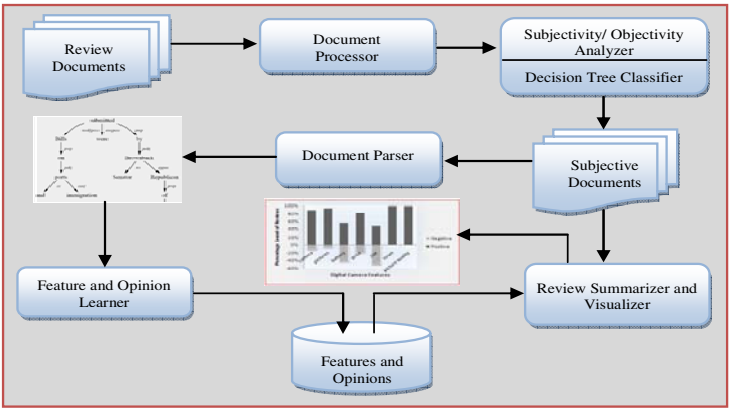

Fig. 1. Architecture of the proposed opinion mining system

Analyzer, *Document Parser*, *Feature and Opinion Learner*, and *Review Summarizer and Visualizer*. The working principles of these components are explained in the following sub-sections.

3.1 Document Processor and Subjectivity/Objectivity Analyzer

Subjective sentences are expressive of the reviewer's sentiment about the product, and objective sentences do not have any direct or obvious bearing on or support of that sentiment [7]. Therefore, the idea of subjectivity analysis is used to retain segments (sentences) of a review that are more subjective in nature and filter out those that are more objective. This increases the system performance both in terms of *efficiency* and *accuracy*. We employ the *Document Processor* which consists of a Markup Language (ML) tag filter, divides an unstructured web document into individual record-size chunks, cleans them by removing ML tags, and presents them as individual unstructured record documents for further processing.

The cleaned documents are converted into numeric-vectors using unigram model for the purpose of subjectivity/objectivity analysis. In document vectors a value represents the likelihood of each word being in a subjective or objective sentence. We have used a corpus of subjective and objective sentences described in [7] for training purpose. The training set is used to get the probability for each word to be subjective or objective. The Decision Tree classifier of Weka[1] is trained to classify the unseen review sentences into subjective and objective classes.

3.2 Document Parser, and Feature and Opinion Learner

The *Document Parser* module uses Stanford parser, which assigns Parts-Of-Speech (POS) tags to every words based on the context in which they appear. The POS information is used to locate different types of information of interest inside text documents. For example, generally noun phrases correspond to product features, adjectives

[1] http://www.cs.waikato.ac.nz/~ml/weka/

represent opinions, and adverbs are used as modifiers to represent the degree of expressiveness of opinions. Since, it is observed that opinion words and product features are not independent of each other rather, each sentence is also converted into dependency tree using the parser. The dependency tree, also known as word-word relationship, encodes the grammatical relations between every pair of words.

The *Feature and Opinion Learner* module is responsible to extract feasible information components from review documents which is analyzed further to identify product features and opinions. It takes the *dependency tree* input and output feasible information components after analyzing noun phrases and the associated adjectives possibly preceded with adverbs. On observation, we found that product features are generally noun phrases and opinions are either only adjectives or adjectives preceded by adverbs. Therefore, we have defined information component as a triplet $<\mathcal{F}, \mathcal{M}, \mathcal{O}>$ where, $\mathcal{F}$ is a noun phrase, $\mathcal{O}$ is adjective possibly representing product feature and $\mathcal{M}$ is adverb that acts as modifier to represent the degree of expressiveness of $\mathcal{O}$. $\mathcal{M}$ is also used to capture negative opinions explicitly expressed in reviews. The information component extraction mechanism is implemented as a rule-based system which analyzes dependency tree to extract information components.

Though a large number of commonly occurring noun and adjective phrases are eliminated due to the design of the information component itself, it is found that further processing is necessary to consolidate the final list of information components and thereby the product features and opinions. During the consolidation process, we take care of two things. In the first stage, since product features are the key noun phrases on which opinions are applied, so a feasible collection of product features is identified using term frequency (*tf*) and inverse document frequency (*idf*). In the second stage of analysis, however, for each product feature the list of all opinions and modifiers is compiled that are used later for polarity determination of the opinion sentences. A partial list of product features, opinions, and modifiers extracted from a corpus of 286 customer reviews on *digital camera* is shown in table 1.

Table 1. A partial list of extracted features, opinions and modifiers for digital camera

Product	Feature	Modifier	Opinion
Digital Camera	picture	not, really, very	beautiful, clear, fantastic, good, great, professional, sharp
	battery	very	decent, excellent, rechargeable, short, long
	price	---	cheap, excellent, good, great, high

3.3 Review Summarizer and Visualizer

In order to generate feature-based summary of review documents, firstly, the polarity of extracted opinions for each feature are classified using Senti-WordNet [8], a lexical resource in which each WordNet synset s is associated to three numerical scores $Obj(s)$, $Pos(s)$ and $Neg(s)$, describing how objective, positive, and negative the terms contained in the synset are. For each feature, the opinion sentences are examined and mapped into one of the *positive* or *negative* class based on the maximum score value of the opinions present in them. In case of presence of multiple features in an opinion

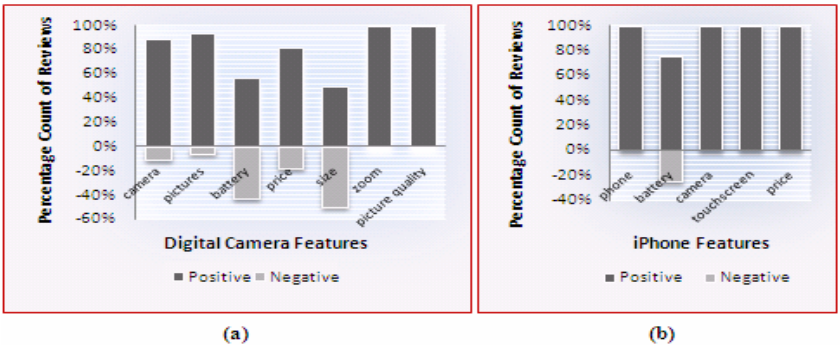

Fig. 2. A feature-based summary generated by the proposed opinion mining system for (a) Digital camera, and (b) iPhone

sentence, the one having highest score value is used to decide its class. Finally, the total number of positive, and negative opinion sentences for each feature is calculated to generate a feature-based review summary which is presented to user in a graphical way as shown in Fig. 2.

4 Evaluation

Since terminology and complex proper names are not found in Dictionaries, an obvious problem of any automatic method for concept extraction is to provide objective performance evaluation. Therefore manual evaluation has been performed to judge the overall performance of the proposed system. From the extraction results, the value of performance measures frequently used for information retrieval tasks - *precision*, *recall*, *F1-measure* and *accuracy* is calculated for each category of experimental data. Table 2 summarizes the performance measure values for our system. The recall value is lower than precision indicating that certain correct feature-opinion pairs could not be recognized by the system correctly. This is justified since most of the reviewers do not follow grammatical rules strictly while writing reviews due to which the parser fails to assign correct POS tag and thereby correct dependency relations between words. However, almost all identified feature-concept pairs are correct, which leaves scope for enhancing our grammar to accommodate more dependency relations.

Table 2. Performance evaluation of feature-opinion extraction process

Product Name		TP	FP	FN	TN	Precision (%)	Recall (%)	F1-measure (%)	Accuracy (%)
Digital Camera	Canon	37	03	27	320	92.50	57.81	71.15	93.46
	Kodak	55	03	73	365	94.83	42.97	59.14	85.02
	Nikon	44	04	63	390	91.67	41.12	56.77	87.85
	Panasonic	32	03	18	155	91.43	64.00	75.29	89.90
iPhone		23	04	14	185	85.19	48.94	62.16	88.14
					Macro-Average	**91.12**	**50.97**	**64.90**	**88.87**

5 Conclusion and Future Work

In this paper, an opinion mining system is proposed to identify product features and opinions from review documents. The proposed method also finds the sentiment polarity of opinion sentences using Senti-WordNet and provides feature-based review summarization and visualization. Presently, we are refining the rule-set to consider more relations to improve the *accuracy* of the system. We are developing a query-answering system to handle opinion-based queries over review documents.

References

1. Liu, B., Hu, M., Cheng, J.: Opinion Observer - Analyzing and Comparing Opinions on the Web. In: Proceedings of the 14th International Conference on World Wide Web (WWW 2005), Japan, pp. 342–351 (2005)
2. Ding, X., Liu, B., Philip, S.Y.: A Holistic Lexicon-Based Approach to Opinion Mining. In: Proceedings of the 1st ACM International Conference on Web Search and Data Mining (WSDM 2008), California, USA, pp. 231–240 (2008)
3. Hu, M., Liu, B.: Mining and Summarizing Customer Reviews. In: Proceedings of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2004), USA, pp. 168–177 (2004)
4. Pang, B., Lee, L., Vaithyanathan, S.: Thumbs up? Sentiment Classification Using Machine Learning Techniques. In: Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP 2002), USA, pp. 79–86 (2002)
5. Popescu, A.M., Etzioni, O.: Extracting Product Features and Opinions from Reviews. In: Proceedings of the 2005 Conference on Empirical Methods in Natural Language Processing (EMNLP 2005), Canada, pp. 339–346 (2005)
6. Turney, P.: Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. In: Proceedings of the 40th Annual Meeting on Association for Computational Linguistics (ACL 2002), Philadelphia, Pennsylvania, pp. 417–424 (2002)
7. Pang, B., Lee, L.: A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts. In: Proceedings of ACL 2004, pp. 271–278 (2004)
8. Esuli, A., Sebastiani, F.: SentiWordNet: A Publicly Available Lexical Resource for Opinion Mining. In: Proceedings of 5th Conference on Language Resources and Evaluation, Genova, Italy, pp. 417–422 (2006)