

MINING CONSUMER OPINIONS FROM THE WEB

Christopher C. Yang^{1,2} and Y. C. Wong²

¹ College of Information Science and Technology, Drexel University, USA

² Department of Systems Engineering and Engineering Management, Chinese University of Hong Kong, Hong Kong

Keywords: Class association rules mining, market intelligence, Web content mining, knowledge management.

Abstract: The Web has provided an excellent platform for business to consumer (B2C) electronic commerce. B2C electronic commerce offers convenience, choice, lower cost and customization to consumers. The electronic shopping platform allows consumers to make intelligent comparison and purchasing decision on consumer products. In addition to comparing product specifications as described on electronic catalogue for better purchasing decision, consumers also hunger for consumer reviews to identify the best products that fit their preferences. For example, a professional photographer would like to identify a camera with lens of high quality and zooming power but a general user may like to find a camera that is cheap, light, and with a large LCD screen. When consumers take consumer reviews as reference, they are interested in both opinion orientation and product features that they are describing. Most of the prior works on consumer opinions mining focus on identifying opinion orientation. Some recent works have started to classify product features but heavily rely on linguistic and natural language processing techniques. However, the writing in consumer reviews is usually less formal and many of them do not conform to the grammatical rules. Therefore, the linguistic and language processing approach is not satisfactory. In this work, we propose a sentiment analysis system to classify product features of consumer reviews by mining class association rules. The experimental result shows that the performance is promising. The content mining approach outperforms the natural language processing approach.

1 INTRODUCTION

Due to the popularity of B2C electronic commerce, large amount of information about consumer products are available on the Web. Such information not only includes product specifications and prices but also consumer reviews. Before making any purchasing decision, consumers usually compare similar products to identify the product with the best specification and lower cost. Infomediaries are available to make intelligent product matching and comparison for multiple e-stores (Bhargava et al., 2000; Menczer et al., 2002; Wong and Yang, 2005; Yang and Wong, 2006). However, such information does not allow consumers to compare the quality of products. Consumers have to rely on other consumers' experience on these products to determine the product that satisfies their expectation. For example, a consumer finds several car models with similar specifications but he is also concern about the reliability and comfort of the model. The information about product features such as reliability

and comfort is not available on the product specification section of electronic catalogues and therefore cannot be compared by infomediaries. Consumers need to rely on the consumer reviews to see if there are any consumer comments about these product features. If many consumers comment that a particular car model is poor in reliability, a consumer will avoid such model although its price is good. As a result, there is a desire of a system that is capable to conduct sentiment classification and analysis automatically and provides a summary of product features comparison.

Currently, there are many Web sites that provide systems for users to submit and search for consumer reviews but these systems are lack of the functionality to compare products by the user-provided sentiment information. For instances, epinion.com, Rateitall.com, and c|net.com are Web systems that collect consumer reviews for various consumer products. These Web sites use a combination of formats to collect consumer comments. In general, the formats include free text, pros, cons, and ratings in n-point scale (n usually

equals 5 or 10). When a consumer searches for the consumer reviews of a particular product, a list of review comments by different users will be given. However, it does not provide any analysis of the comments nor any summary of the comparison on the product features. There are a number of research efforts focusing on determining the opinion orientation of consumer reviews. The recent work extends the prior effort to classify the product features that are described in the opinion sentence. Most of them rely on linguistic and natural language processing techniques to determine the product feature that are described in a sentence. The performance is good if the sentence is correctly tagged by the natural language processor. However, the natural language processing rules are not able to tag a sentence if it does not conform to the linguistic and grammatical rules of a language. Writing in the Web is not as rigorous as writing in a formal document such as business report or journal article. Therefore, the sentences appearing on Web consumer reviews have many grammatical errors and are not necessary complete sentences. The performance of the linguistic and natural language approach is not satisfactory. In the next section, we provide a literature review of related work.

1.1 Literature Review

The related work in the literature includes sentiment classification and sentiment analysis. Sentiment classification determines the sentiment orientation (either positive or negative) of an opinion text. It only captures the general opinion orientation of a consumer product. However, it doesn't mean that all product features of a product have the same orientation as the general orientation. A camera that is good in general does not mean that its battery life must be long. It may only mean that most of its product features are good and therefore the general orientation is good. Unfortunately, some consumers have expectation on a specific product feature. Without satisfying such specific expectation, the purchasing decision cannot be made. Sentiment analysis takes a further step to classify the product features that an opinion sentence is describing. For examples, product features of a camera may include, lens, battery, usability, photo quality, price, etc.

1.1.1 Sentiment Classification

Sentiment classification can be considered as a type of text classification. In general, text classification techniques classify a collection of documents into a

number of categories. Each category has a topic. For examples, we may classify news articles into local, international, business, technology, sports, entertainment, science, etc. Sentiment classification classifies documents into positive or negative categories on the basis of the sentiment rather than the topics expressed in the document.

The most important component in sentiment classification is learning a sentiment classification model from a set of pre-classified documents. Different features for model learning have been explored in the previous work. Weibe et al. (Weibe et al. 1999) and Hatzivassiloglou and Wiebe (Hatzivassiloglou and Wiebe, 2000) proposed to identify nouns and adjectives that are indicative of positive or negative opinions. Das and Chen (Das and Chen, 2001) used a manually crafted lexicon in conjunction with scoring methods to classify messages on on-line stock message board. Turney (Turney, 2002) used mutual information between term phrases and the words "excellent" and "poor" to identify words of opinions for sentiment classification. Pang et al. (Pang et al. 2002) and Dave et al. Wei et al. (Wei et al., 2006) employed two comprehensive lists of positive and negative words from the General Inquirer (www.wjh.harvard.edu/~inquirer/) as features.

1.1.2 Sentiment Analysis

Hu and Liu (Hu and Liu, 2004) and Liu et al. (Liu et al. 2005) used association mining to extract the explicit product features in review comments. The product features that are described implicitly will not be considered. The sentences are first tagged and parsed by NLProcessor linguistic processor (www.infogistics.com). The frequent product features will then be extracted. A set of opinion words expanded from a set of seed adjectives using the WordNet will be adopted to determine the opinion orientation. The opinion words are also used to identify infrequent product features. In this approach, sentences that do not include any product feature words and opinion words are not considered. These sentences are not opinion sentences according to their definition although the orientation of a product features are described implicitly. This approach achieves high performance for those sentences that are correctly tagged by NLProcessor. However, many sentences with grammatical errors or incomplete sentences are not able to be tagged correctly. Extensive manual effort is required to adjust the tagging before further sentiment analysis processes.

In this work, we propose the class association rules mining approach to identify the associations between keywords and product feature classes. In this approach, it does not suffer from the unsatisfactory performance of natural language processing due to the informal writing on the Web.

2 CLASS ASSOCIATION RULES

We model the problem of the sentiment analysis of consumer products as class association rule mining. Let $R = \{r_1, r_2, \dots\}$ be a set of consumer reviews of a product. In each consumer review r_i , there are a number of entities, such as the reviewer's id, post date, rating and sentences in free text, pros, and cons areas. Among all these entities, we take the sentences that describe the product in free text, pros and cons areas for analysis. Let $S_i = \{s_{i1}, s_{i2}, \dots\}$ be the set of sentences of r_i . For each product, we have a set of predefined product features, $F = \{f_1, f_2, \dots\}$. We assume that there is a set of keywords associated with each product feature f_j . Let $w_j = \{w_{j1}, w_{j2}, \dots\}$ be the set of keywords associated with f_j . The goal of class association rules mining is extracting the associations between keywords and product feature classes from a set of pre-classified sample sentences collected from the consumer review repository. These associations will then be utilized to classify sentences based on the keywords available in the sentences.

A class association rule is defined as

$$w_{jk} \rightarrow f_j \text{ where } w_{jk} \in w_j \text{ and } f_j \in F$$

Each sentence in a consumer review is represented by a vector which is composed of keywords in the sentence. In general, a consumer review includes sentences that describe the positive and negative opinions of the overall product and the specific product features of the product. However, there are also sentences that do not describe anything about product features but only telling a story about how a consumer uses the product. For example, "I take my new Cannon EOS to a trip in London. I have taken many pictures in downtown." does not describe any product features but only a story. Some sentences may describe more than one product feature. For example, "a bit pricey but great photos" describes two product features, price and photo quality. As a result, by applying the class association rules, a sentence in S_i may not be classified into any product feature class, f_j , but only to a class, NA . In this case, no assignment of class is made for the sentence. A sentence in S_i may also be classified into one or

more than one product feature class f_i . In this case, one assignment or multiple assignments are made for a sentence. After assigning sentences into the product feature classes, a number of sentences are found in each product feature class where these product classes are not mutually exclusive. Some classes may have more number of sentences but different percentages of positive sentences and negative sentences. A summary of comparison across several products can be made on the percentage of positive and negative opinions on each product feature and the frequency of opinion expression of each product feature. An illustration is made on Table 1.

Table 1: An illustration of a summary of comparison by sentiment analysis.

	Product A	Product B	Product C
Product Feature 1	Pros: 80% Freq: 52	Pros: 71% Freq: 29	Pros: 61% Freq: 38
Product Feature 2	Pros: 87% Freq: 50	Pros: 95% Freq: 69	Pros: 72% Freq: 66
Product Feature 3	Pros: 96% Freq: 58	Pros: 50% Freq: 35	Pros: 94% Freq: 40
Product Feature 4	Pros: 80% Freq: 4	Pros: 75% Freq: 3	Pros: 75% Freq: 3
Overall	Pros: 88% Freq: 164	Pros: 72% Freq: 109	Pros: 74% Freq: 147

In the above illustration, it shows that Product A is good in all product features. That means Product A is more preferable than Product B and Product C. However, Product Feature 4 is not frequently discussed in consumer reviews. Therefore, there may not be sufficient data to confirm that Product Feature 4 is necessary good for Product A and indeed for Product B and Product C as well. If Product Feature 2 is the major concern of a consumer but he does not concern about other product features, he may choose Product B instead because it has the highest rating in terms of Product Feature 2. The summary produced by sentiment analysis allows consumer to compare all or individual product features across several products.

2.1 Class Association Rules Mining

Given a set of pre-classified sample sentences, the class association rules mining extracts a set of class association rules, $w_{jk} \rightarrow f_j$. Each sample sentence is tagged with associated product features manually. Some sentences do not have any tags because they are not associated with any product features. Using these tagged sentences, a class associated rule with a keyword or key phrase, w , and a product feature, f , is

extracted if they satisfy the minimum support and minimum confidence.

Support:

$$\text{Support}(w, f) = \text{Freq}(w \text{ and } f) / N$$

Confidence:

$$\text{Confidence}(w, f) = \text{Freq}(w \text{ and } f) / \text{Freq}(w)$$

($\text{Freq}(x)$ is the frequency of x , N is the total number of sample sentences.)

If $\text{Support}(w, f) \geq \text{min_Support}$ and $\text{Confidence}(w, f) \geq \text{min_Confidence}$, then we deduce $w \rightarrow f$.

min_Support and min_Confidence are two parameters in class association rules mining that affect the rules mining result. Tuning of these parameters is usually required to optimize the rules mining performance.

2.2 Firing Multiple Rules and Conflict Resolution

When a sentence describes multiple product features, multiple rules are matched with the sentence and will be fired. In this case, multiple product features will be assigned. However, the matched rules may have conflicts. For example, the word in Rule 1 is part of the phrase in Rule 2. If we are firing both rules when a sentence has the phrase in Rule 2, the sentence will be assigned with two product features according to Rule 1 and Rule 2. The following two rules for a consumer product, camera, are illustrations of such conflict situation. Assuming there are two product features, *Flash* and *Memory*, which are important features that consumers are interested in.

Rule 1: *flash* \rightarrow *Flash*

Rule 2: *flash card* \rightarrow *Memory*

“*flash*” in Rule 1 is part of “*flash card*” in Rule 2. When we say *flash*, it discusses the flash light of a camera. However, when we say *flash card*, it describes the memory card in a digital camera. Without conflict resolution, both Rule 1 and Rule 2 will be fired for a sentence “The new flash card stores images very fast and you won’t miss any shot.” As a result, the sentence will be assigned with two product features, *Flash* and *Memory*. However, the sentence only describes *Memory* but not *Flash*.

We develop a conflict resolution mechanism to resolve the conflict as described. Since a longer phrase describe more specific meaning than a single word or a shorter phrase, the semantic carried by a longer phrase is more reliable than a shorter phrase in determining the product feature that it refers to. For example, “Drexel University” describes a

particular university but “University” describes a university in general. The conflict resolution mechanism compares the words or phrases in two rules. If the word or phrase in a rule is part of the word or phrase in another rule, the rule with a shorter phrase will be removed.

The following algorithm describes the procedures in assigning product features to sentences in a consumer reviewer.

For every sentence in a consumer reviewer

 Identify the matched rules for the sentence

 If no rule can be matched

 Assign the sentence to the class *NA*

 If only 1 rule is matched

 Fire the rule and assign the corresponding product feature to the sentence

 Else

 Active the conflict resolution to remove rules for resolving conflicts

 Fire the remaining rules and assign the corresponding product features to the sentence

3 EXPERIMENT

We have conducted an experiment to evaluate the performance of the proposed class association rules mining approach for sentiment analysis of consumer product reviews.

3.1 Data Collection

We have collected consumer reviews on digital cameras from Amazon.com. In the data collected, we have 3000 sentences from 214 consumer reviews on 6 digital camera models. The six digital camera models are:

- Canon EOS 20D 8.2MP Digital SLR Camera,
- Nikon D70 Digital SLR Camera Kit,
- Canon Powershot SD300 4MP Digital Elph Camera with 3x Optical Zoom,
- Sony Cybershot DSCP200 7.2MP Digital Camera 3x Optical Zoom,
- Canon Powershot S2 IS 5MP Digital Camera with 12x Optical Image Stabilized Zoom
- Canon Powershot A95 5MP Digital Camera with 3x Optical Zoom

We have tagged the sentences with their product features. 56% of the sentences are tagged with *NA*, which means no product feature is identified. 44% of the sentences are tagged with one or more product

features. Among those sentenced tagged with one or more product features, 87% has one product feature, 12% has two product features, 0.9% has three product features and 0.1% has four product features.

Nine product features are identified in the consumer reviews of six digital camera models. They are battery, flash, image quality, lens, memory, price, usability, and video.

3.2 Experiment Metrics

We use precision (P) and recall (R) to measure the effectiveness of our proposed opinion sentence identification method for the focused sentiment analysis. For any product feature f_j ,

$$P_{f_j} = \frac{TP_{f_j}}{TP_{f_j} + FP_{f_j}} \text{ and}$$

$$R_{f_j} = \frac{TP_{f_j}}{TP_{f_j} + FN_{f_j}}$$

where TP_{f_j} is the number of true opinion sentences of the product feature f_j that are correctly identified by our proposed method, FP_{f_j} is the number of non-opinion sentences of the product feature f_j that are incorrectly identified by our proposed method as the opinion sentences of f_j , and FN_{f_j} is the number of true opinion sentences of the product feature f_j that are missed by our proposed method.

The precision measures how precise are the sentences that we have identified with a product feature f_j . The recall measures how many sentences our proposed method has identified with a product feature f_j among those sentences with f_j in the whole collection.

To assess the overall performance across all product features, we adopt both micro average and macro average:

$$\text{Micro-precision} = \frac{\sum_{f_j \in F} TP_{f_j}}{\sum_{f_j \in F} TP_{f_j} + \sum_{f_j \in F} FP_{f_j}}$$

$$\text{Micro-recall} = \frac{\sum_{f_j \in F} TP_{f_j}}{\sum_{f_j \in F} TP_{f_j} + \sum_{f_j \in F} FN_{f_j}}$$

$$\text{Macro-precision} = \frac{\sum_{f_j \in F} P_{f_j}}{|F|}$$

$$\text{Macro-recall} = \frac{\sum_{f_j \in F} R_{f_j}}{|F|}$$

3.3 Experimental Result

We have conducted a 5-fold cross validation experiment using $min_support = 0.006$ and $min_confidence = 0.5$. A set of tagged sample sentences are used to learn the class association rules. The extracted class association rules are then used to determine the product features described in a set of testing sentences without any tagging. Such 5-folded learning and testing processes are conducted on the collected data as described in Section 3.1 to produce the experimental results. Table 2 shows the experimental result.

Table 2: Experimental Result.

Product Feature	Precision	Recall
<i>Lens</i>	80.42%	90.37%
<i>Memory</i>	72.92%	91.56%
<i>Flash</i>	70.04%	90.86%
<i>Price</i>	78.98%	73.92%
<i>Image Quality</i>	55.25%	45.65%
<i>Battery</i>	88.24%	92.90%
<i>Screen</i>	77.05%	95.82%
<i>Video</i>	79.50%	88.05%
<i>Usability</i>	65.64%	67.73%
Micro average	72.77%	76.12%
Macro average	74.23%	81.87%

Some product features achieve higher precision and recall while some product features suffer in relatively low precision and recall. For example, *Battery* achieves a precision of 88.24% and a recall of 92.9% while *Image Quality* only achieves a precision of 55.25% and a recall of 45.65%. It is mainly due to the usage of words in describing a product feature and how precise the definition of a product feature is. For the product feature *battery*, the words describing battery are very specific, e.g. charger, charging, Li-ion, battery life. Other product features such as *Memory* also have very specific descriptive words such as memory card, flash card, SD card, MB, GB, etc. On the other hand, for the product feature *Image Quality*, consumers use a wide variety of words to describe the quality of image (e.g. blurry, color, nice picture etc.) and the words for describing *Image Quality* may be used in other sentences that do not describe any product features. In some cases, even human judges are also difficult to determine if the sentence is describing the quality of image when a consumer is telling a story about the images that he has taken. The classification of the product feature is relatively vague comparing with other product features.

In general, we find the proposed sentiment analysis technique is promising. The micro

precision and recall are 72.77% and 76.12%. The macro precision and recall are 74.23% and 81.87%.

4 CONCLUSIONS

As B2C electronic commerce is so popular nowadays, it is convenience to shop online. However, consumers are expecting more than just convenience, choice, and lower price as offered by B2C electronic commerce. Consumers desire to have an intelligent Web information system to support their purchasing decisions. In such system, consumer reviews can be compared in terms of their product features so that consumers are able to identify the best consumer products that they want.

In this work, we propose a Web content mining approach using class association rules mining to overcome the problems that exist in the traditional linguistic and natural language processing approach for sentiment analysis. The writing on consumer reviews is usually informal and contains a lot of grammatical errors. As a result, many sentences cannot be correctly parsed and further processed to determine the product features that they are describing. In our class association rules mining approach, sentences are not required to be parsed by natural language processor. Based on the usage of words and their frequency, we determine the relationships between words (or phrases) and product features classes. Using the learned rules, we determine the product features described in an opinion sentence.

Our experimental result shows that it achieves a promising performance. The performance is especially good for the product features that are specific and clear. However, it still suffers in relatively lower precision and recall when the product features are not as well-defined.

In the future, we shall investigate other statistical techniques to support the selection of words so that words with higher discriminating power can be identified to produce better performance in sentiment analysis.

REFERENCES

- Bhargava, H. K., Choudhary, V., Krishnan, R., 2000. Pricing and Product Design: Intermediary Strategies in an Electronic Market. In *International Journal of Electronic Commerce*, Vol. 5, No. 1, pp.37-56.
- Das, S., Chen, M., 2001. Yahoo! for Amazon: Extracting Market Sentiment from Stock Message Boards. In *Proceedings of the 8th Asia Pacific Finance Association (APFA) Annual Meeting*.
- Hatzivassiloglou, V., Wiebe, J., 2000. Effects of Adjective Orientation and Gradability on Sentence Subjectivity. In *Proceedings of 18th International Conference on Computational Linguistics (COLING)*, Saarbrücken, Germany, pp.299-305.
- Hu, M., Liu, B., 2004. Mining and Summarizing Customer Reviews. In *Proceedings of the 10th ACM Conference on Knowledge Discovery and Data Mining (KDD)*, Seattle, WA, pp.168-177.
- Liu, B., Hu, M., Cheng, J., 2005. Opinion Observer: Analyzing and Comparing Opinions on the Web. In *Proceedings of the 14th International Conference on World Wide Web (WWW'05)*, Chiba, Japan, pp.342-351.
- Menczer, F., Street, W.N., Monge, A. E., Adaptive Assistants for Customized E-shopping. In *IEEE Intelligent Systems*, Vol. 17, No. 6, pp. 12-19.
- Pang, B., Lee, L., Vaithyanathan, S., 2002. Thumbs Up? Sentiment Classification Using Machine Learning Techniques. In *Proceedings of 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP 2002)*, pp.79-86.
- Turney, P., 2002. Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. In *Proceedings of the 40th Conference on Association for Computational Linguistics (ACL)*, Philadelphia, PA, pp.417-424.
- Wiebe, J., Bruce, R., O'Hara, T., 1999. Development and Use of a Gold Standard Data Set for Subjectivity Classifications. In *Proceedings of the 37th Conference on Association for Computational Linguistics (ACL)*, College Park, MD, pp.246-253.
- Wei, C., Yang, C. S., Huang, C. N., 2006. Turning Online Product Reviews to Customer Knowledge: A Semantic-based Sentiment Classification Approach. In *Proceedings of 10th Pacific Asia Conference on Information Systems (PACIS)*, Kuala Lumpur, Malaysia.
- Wong, R., Yang, C. C., 2005. Collaborative Infomediary for Financial News. In *Proceedings of the Fourth Workshop on e-Business (WEB 2005)*, Las Vegas, NV, December 2005.
- Yang, C. C., Wong, R., 2006. Measuring Success Factors of E-Commerce Infomediary. In *Proceedings of the Pacific Asia Conference on Information Systems (PACIS)*, Kuala Lumpur, Malaysia, July 2006.