Volume 6, No. 3, May 2015 (Special Issue)
ISSN No. 0976-5697
International Journal of Advanced Research in Computer Science
RESEARCH PAPER
Available Online at www.ijarcs.info

Mining Online Customer Reviews for Product Feature-Based Ranking

K.Parish Venkata Kumar
Asst. Professor
Dept.of Computer Applications
V.R.S.E.
VIJAYAWADA, INDIA
parish4u@gmail.com

G. Sai Chaitanya Kumar
Associate..Professor
Dept.of.CSE
PPDCE &T
VIJAYAWADA, INDIA
sai.dmca@gmail.com

M.Aruna
Asst.Professor
Dept.of.CSE
NRI Institute of Technology
VIJAYAWADA, INDIA
marunamtech@gmail.com

B. Srinivas
Asst.Professor
Dept.of.Computer Applications
V.R.S.E.C
VIJAYAWADA, INDIA
sri.happydays@gmail.com

Abstract: Recent trends have indicated that large numbers of customers are switching to online shopping. Online customer reviews are an unbiased indicator of the quality of a product. However, it is difficult for users to read all reviews and perform a fair comparison. We describe a methodology and algorithm to rank products based on their features using customer reviews. First, we manually define a set of product features that are of interest to the customers. We then identify subjective and comparative sentences in reviews using text mining techniques. Using these, we construct a feature-specific product graph that reflects the relative quality of products. By mining this graph using a page-rank like algorithm (pRank), we are able to rank products. We implement our ranking methodology on two popular product categories (Digital Camera and Television) using customer reviews from Amazon.com. We believe our ranking methodology is useful for customers who are interested in specific product features, since it summarizes the opinions and experiences of thousands of customers.

Keywords: Text mining, page-rank, quality.

I. INTRODUCTION

Increasingly large numbers of customers are choosing online shopping because of its convenience reliability, and cost. As the number of products being sold online increases, it is becoming increasingly difficult for customers to make purchasing decisions based on only pictures and short product descriptions. On the other hand, customer reviews, particularly the text describing the features, comparisons and experiences of using a particular product provide a rich source of information to compare products and make purchasing decisions. Online retailers like Amazon.com allow customers to add reviews of products they have purchased. These reviews have be- come a diverse and reliable source to aid other customers. Traditionally, many customers have used expert rankings which rate limited a number of products. Existing auto- mated ranking mechanisms typically rank products based on their overall quality. However, a product usually has multiple product features, each of which plays a different role. Different customers may be interested in different features of a product,and their preferences may vary accordingly. In this work, we present a feature-based product ranking technique that mines thousands of customer reviews. We first identify product features within a product category and analyze their frequencies and relative usage. For each feature, we identify subjective and comparative sentences in reviews. We then assign sentiment orientations to these sentences. By using the information obtained from customer reviews, we model the relationships among products by constructing a weighted and directed graph. We then mine this graph to determine the relative quality of products. Experiments on Digital Camera and Television reviews from real-world data on Amazon.com are presented to demonstrate the results of the proposed techniques.

This work is based on the our proposed ranking scheme, where products are ranked based on their overall quality. A product feature is defined as an attribute of a product that is of interest to customers. Even though an overall ranking is an important measure, different product features are important to different customers based on their usage patterns and requirements. And different products may be designed and rank differently based on the feature of interest. For instance, a digital camera that is ranked highly overall may have lessthan-stellar battery life. Thus, both overall ranking and more detailed product feature based ranking are important. In this work, we propose an algorithm that uses customer review text, mines tens of thousands of reviews, and provides ranking of products based on product features. For each product category, product features are defined and extracted in the data preprocessing step. Note that for most products, there is a standard set of features which are considered important and normally are provided with product descriptions. We then label each sentence with the product features described in it. We then identify four different types of sentences in customer reviews that are useful in determining a products rank: positive subjective, negative subjective, positive comparative, negative comparative. Subjective sentences are those sentences in which the reviewer expresses positive or negative sentiments about the product being reviewed. Comparative sentences contain information comparing competing products in terms of features, price, reliability etc.

II. PRODUCT RANKING METHOD

Product Feature Identification Product features are attributes that provide added functionality to products. In the

CONFERENCE PAPER
4th National Conference on Recent Trends in Information Technology 2015 on 25/03/2015
Organized by Dept. of IT, Prasad V. Potluri Siddhartha Institute of Technology, Kanuru, Vijayawada-7 (A.P.) India

product description, manufacturers tend to highlight their product features from different (often contradicting) perspectives. The combination of features available in a product influences the purchasing decision of a customer. Some consumers typically seek an optimized combination of features to satisfy their needs, while others may focus on a single feature. There has been some research on automatically identifying the different features of a product that customers are concerned about. There has been earlier research on identifying product features and feature sentences, which is not the focus of this work. We assume that product features associated with a product domain are given. We manually gathered product feature sets for two product categories: Digital Cameras and Television based on the official consumer reports. This is a one-time pre-processing overhead that can be done relatively easily by someone being familiar with a product domain.

The use of synonyms is motivated by the fact that customers use different words/spelling variants to describe product features. To test the effectiveness of our feature-finding mechanism, we randomly picked 1000 review sentences from our review pool and manually labeled each sentence with the product features This small dataset was used to evaluate the performance of our feature-finding strategy. The precision and the recall of the keyword strategy for digital camera data were 0.853 and 0.807 respectively. Therefore, we are able to find a major portion of feature sentences using this simple yet effective strategy.

In our review dataset containing 1, 516, 001 sentences, we observed that around 16% of sentences describe one or more of these features. To tag each sentence with features, we use a simple strategy: if the sentence contains one of the words/phrases in the synonym set for a product feature, we mark it as describing the feature. Since we have defined these features manually, we describe them in greater detail. For Digital Cameras, ash is an important feature for indoor and low-light photography. Battery life is a sought-after feature that details the kind of batteries used. Focus talks about auto-focus or manual focus capabilities. Lens is a critical factor for professional photographers purchasing high-end cameras. The Optical feature encompasses digital zoom and optical zoom. LCD represents the digital display/screen that lets a user see how a photo will look like. Resolution refers to the sharpness, or detail, of a picture. Burst is used to describe the rapid fire and continuous shooting capabilities. Memory determines the number of pictures that can be taken and Compression determines how _le size of a photo is shrunk. For the Television segment, the Sound feature is useful for users interested in audio quality (some TVs come with an extra set of speakers to create surround sound). Reflectivity/Anti-glare is important for the viewing experience. Size represents the size, height, weight of a television screen. Connections mean the number and type of input ports available for hooking up devices to the television. The richness/quality of the images displayed are described in Picture quality. Users are also interested in the remote control device available with the television. Resolution refers to the number of pixels or lines displayed on the screen. Adjustment is the ability/mode that expands or compresses an image to fill the screen better. Picture-in-picture(PIP) feature allows a user to watch two channels at once. Film- Mode/Cine Motion improves the movie watching experience, which may be important to some users.

II.1. Identifying Comparative Sentences

There has been some earlier research regarding identification of comparative sentences in text. These techniques use keyword comparison (KW contains 126 words, some of which are explicit (outperform, exceed, compare, superior, etc.) and others are implicit (prefer, choose, like, etc.), sentence semantics, and sentence structure to identify comparative sentences. To identify part-of-speech tags, CRFTagger , a java-based conditional random field part-of-speech (POS) tagger for English is employed. We build on these techniques and use the following rules for identifying comparative sentences:

-> Check if the sentence contains any comparative key-words in KW;

-> Recognize any words with POS tags € JJR (comparative adjective), RBR (comparative
ad- verb), JJS (superlative adjective), RBS (superlative adverb);

-> Scan if any predefined structural patterns are present in the sentence (as <word>
as, the same as, similar to, etc.).

Note that not all sentences satisfying these rules are comparative sentences in terms of product comparison. For example, the sentence I bought this camera for my son because he got a higher grade in his second statistical exam." does not show any comparative meanings or implications over other camera products. Therefore, we propose a more refined technique to find comparative sentences specifically related to product comparisons in our previous work. We use a dynamic program- ming technique (longest common subsequence) to identify product-product comparison pairs in a comparative sentence. We use only comparative sentences which contain at least one product name which is di_erent from the product the sentence is describing while building our ranking model. In previous work, we have shown that we get a precision of 82% and a recall of 80% approximately.

II.2 Constructing the Product Graph

We use the subjective and comparative sentences found to construct a directed and weighted graph that can be mined to reveal the relative quality of products. The graph is defined as follows: G = {V,E} where

-> V is the set of nodes, V = {pi| each product represents a node, 0 < i < n}

-> E is the set of directed edges. An arc e = (pi, pj ,wij) is considered to be directed
from pi to pj with a weight wij . E = {ek = (pi, pj, wij)| wij is the weight of the
edge eij, 0<_ i, j < n, 0<_ k < m)

where n is the number of products, m is the number of edges.

Consider a comparative sentence in the reviews for a product Pi. If this sentence compares
product Pi with product Pj , we add an directed edge from Pj to Pi. The second step is
to assign a weight to this edge. A comparative sentence occurring in the reviews for product
Pi and comparing it with product Pj is considered a positive comparative (PC(Pi, Pj)) if it
implies that Pi is better than Pj . If it implies that Pi is worse that Pj , it is considered a
negative comparative (NC(Pi, Pj)). For each edge(Pj ; Pi), we count the number of positive
(PC) and negative (NC) comparative sentences associated with the pair (Pi; Pj) respectively.
We assign the score based on PC;NC as the weight of the edge linking Pj to Pi. The last step is to assign weights for nodes. For a node Pi, we use the score derived from the number of positive(PS) and negative (NS) subjective sentences (PS;NS) as its weight. In our algorithm, we take the degree of confidence of each review into account when calculating the weight of nodes and edges. In addition, we also assume that the number of reviews also plays an important role. Based on these assumptions, the score could be calculated as follows.

$$\mathbf{S_i} = \sum_{k=1}^{m} C_k \log\left(\frac{Pos_k}{Neg_k} + 1\right),$$

where:
-> m is the number of reviews for a specific product Pi,
-> Ck is the degree of confidence for the kth review,
-> Posk,Negk are the number of positive and negative sentences within the kth review.
They are corresponding to PSk,NSk, or PCk,NCk for node weight and edge weight
Respectively

$$C_k = \begin{cases} \frac{\#\, helpfulvotes_k}{\sum_{t=1}^{T} \#\, totalvotes_t} * \frac{\#\, helpfulvotes_k}{\#\, totalvotes_k} & \#\, totalvotes_k \neq 0; \\ 1 & \text{otherwise.} \end{cases}$$

III. pRank: Product Ranking Algorithm

We propose a product ranking algorithm based on the concept of the PageRank algorithm[?]
to evaluate the relative importance of each product. In our ranking algorithm: pRank, a node (product) has a higher importance if it is pointed (favorably compared) to from relatively important nodes. We not only consider the relative importance among products, but also take the importance of the product itself into account. This means that the node weight is also crucial to the ranking, in addition to the edge weights. In practice, for some product categories, as the number of comparative sentences is relatively small and may cover a small portion of products, the graph constructed from those can be sparse. The node weights generated from subjective sentences play a more important role in ranking score calculation. The termination of the pRank algorithm is dependent on the damping factor.

$$\text{pRank}(P) = [(1-d) + d * \textstyle\sum_{i=1}^{n} 1_{\{P_i,P\}} * pRank(P_i) * C_e(P_i)] * C_v(P),$$

where:

->pRank(P) is the product ranking of product P;
->pRank(Pi) is the product ranking of product Pi and n is the number of incoming links on product P;
->1{Pi,P } is an indicator function, s.t.

$$1_{\{P_i,P\}} = \begin{cases} 1 & \text{if there is a link from } P_i \text{ to } P; \\ 0 & \text{otherwise.} \end{cases}$$

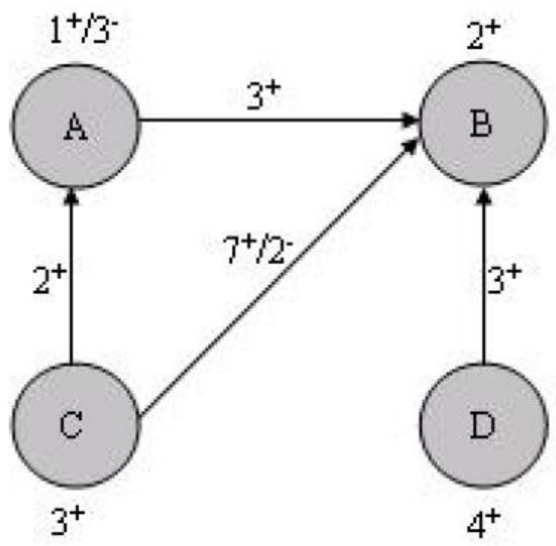

Figure 1.1. A simple ranking example with a product graph Gf having 4 products regarding a specific feature f (+ means positive and □ means negative).

- $C_e(P_i) = \frac{W_e(P_i,P)}{\sum_{j=1}^{m} W_e(P_i,P_j)}$, where m is the number of outbound links on product P_i, P_j are the nodes pointed to from P_i and $W_e(P_i, P_j)$ is the weight of the edge (P_i, P_j). It is the edge weight contributor to the ranking of product P;
- $C_v(P) = \frac{W_v(P,P)}{\sum_{t=1}^{n} W_v(P_t,P_t)}$. It is the node weight contributor to the ranking of product P.

Let us illustrate the ranking process using a simple example. We have four products (A, B, C,D) which we wish to rank according to product feature f. The numbers of positive/negative,
subjective/comparative sentences labeled with feature f are listed below.

PSf (A) = 1; PSf (B) = 2; PSf (C) = 3; PSf (D) = 4NSf (A) = 3;
PCf (B,A) = 3; PCf (B,C) = 7PCf (B,D) = 3; PCf (A,C) = 2;NCf (B;C) = 2

Based on these sentence statistics, we could build a product graph G (see Fig. 1.1). Edge weights are determined by comparative sentences, and node weights are determined by subjective sentences. Since the reviews of product B have 4 positive comparative sentences mentioning product C and 2 negative comparative sentences mentioning C from different reviews, there is an edge from C to B with weight 3:1239 based on our scoring strategy. It must be mentioned that to prevent edges with in_nite length (when the number of negative comparative sentences is 0), we set the minimum value of the denominator to 1 while computing edge weights.
By using our algorithm, we get the ranking score for each node shown in the Table 1.1.The ranking order (the smaller, the product better) for this graph is B -> D-> C-> A.From the graph, we clearly see that A, C, D are worse than B because all of them have edges pointing to B. D has more positive subjective sentences than A, C and their comparative weights with B are approximately equal. C has a better ranking than A because (i) two sentences say A is better than C and (ii) reviews for A contain 1 positive/3 negative subjective sentences while reviews for C contain 3 positive subjective sentences

Table 1.1. Product ranking results for Gf in Fig. 1.1

Rank Order	Ranking Score	Vertex
1	0.820731	B
2	0.072917	D
3	0.053571	C
4	0.052781	A

IV. Experimental Results On Online Customer Reviews

In this section we evaluate the performance of our ranking algorithm. We conduct our experiments on customer reviews from two different product categories (Digital Camera and Televisions) from Amazon.com. Further details about the datasets and the APIs used to generate this data can be found at Amazon.com2 and BrowseNodes.com3. The Digital Camera dataset contains 83005 reviews (for 1350 products) and Television dataset contains 24495 reviews (for 760 products) collected by August, 2009. Table 1.2 show the relevant statistics for these two datasets: total number of sentences, frequency of occurrence of different product features, number of subjective and comparative sentences and their sentiment orientations. To evaluate our ranking algorithm, we first perform product ranking based on the overall quality. To determine the overall rank of a product, we include all comparative and subjective sentences in our database while constructing he product graph.

There is no filtering done for product features. We then mine this overall graph G using the ranking algorithm described above. To evaluate the effectiveness of this ranking strategy, we compare our results with a ranking performed by domain experts. The results indicate that our product ranking strategy achieves significant agreement with evaluations done by subject experts with several years of experience and insight in their respective fields. Approximately, an average overlapping probability of 62% could be achieved for different price bins for cameras and televisions. In this work, we focus on the feature-specific ranking obtained by mining the individual product graphs generated for each product feature. Intuitively, the feature-specific ranking should not be dramatically different from the overall ranking. If we have chosen a relevant set of product features that customers are interested

Table 1.2. Breakdown of subjective/comparative sentences (digital camera).

Feature/Overall	#Sentences	#Subjective Sentences		#Comparative Sentences	
		Positive	Negative	Positive	Negative
Flash	48378	10045	8202	1358	514
Battery	42461	4838	6439	1030	533
Focus	42393	7306	7241	1389	720
Lens	36371	4678	5313	1055	437
Optical	28658	3771	3196	842	338
Lcd	25874	4357	3587	755	216
Resolution	14992	1768	1647	579	227
Burst	14362	2925	2726	523	189
Memory	10794	1225	1652	365	143
Compression	1780	225	236	78	29
Digital Camera	1469940	71565	97349	16246	10890

in, then the top-ranked products in these lists should not rank badly in the overall list. However, it is quite likely that there are significant differences in the ranking order of these products, especially at the top. To clarify this intuition, we give the following example: If a product ranks in the top 5 products according to feature f (e.g. lens), then the probability that it ranks in the bottom 5 products overall should be very low. Similarly, if a product
has high overall rank, then it should rank highly according to some features

V. Conclusion

CONFERENCE PAPER
4th National Conference on Recent Trends in Information Technology 2015 on 25/03/2015
Organized by Dept. of IT, Prasad V. Potluri Siddhartha Institute of Technology, Kanuru, Vijayawada-7 (A.P.) India

To conclude, I studied social sentiment identification using rule-based, learning-based, and social network-based methods. By incorporating social sentiments, we analyze large amounts of online customer reviews to rank products in order to help people make informed decisions. Furtherly, I also studied measuring social brand reputation through building a probabilistic graph among social users and social brands. A parallelized block-based MCMC technique was proposed to infer brand reputation based on observed sentiments of comments made by users on brands. In addition, a predictive user model was also build to improve online social advertisements. However, there are still many interesting research questions that I would like to pursue over the coming career years. Some of them are motivated by the projects I described above, others are motivated by the collaborations with Faculty or researchers in the communication school, business school, and eBay research lab.

VI. References

[1] Community evolution mining in dynamic social networks. Procedia - Social and Behav- ioral Sciences,

[2] http://www.marketingmag.com.au/news/social-ad-revenue-to-hit-11-billion-in-us-by-2017-387032013.

[3] Charu C. Aggarwal and Philip S. Yu. Outlier detection with uncertain data. In SDM'08, pages 483{493, 2008.}

[4] Amr Ahmed, Yucheng Low, Mohamed Aly, Vanja Josifovski, and Alexander J. Smola. Scalable distributed inference of dynamic user interests for behavioral targeting. In Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining, KDD '11, pages 114{122, New York, NY, USA, 2011. ACM.

[5] Sitaram Asur and Bernardo A. Huberman. Predicting the future with social media. In Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology - Volume 01, WI-IAT '10, pages 492{499, Washington, DC, USA, 2010. IEEE Computer Society

[6] J. Liu, Y. Cao, C.-Y. Lin, Y. Huang, and M. Zhou. Low-quality product review detection in opinion summarization. In EMNLP-CoNLL(Poster), pages 334:342, 2007.

[7] Jingjing Liu and Stephanie Sene_. Review sentiment scoring via a parse-and-paraphrase paradigm. In Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 1 - Volume 1, EMNLP '09, pages 161{169, Stroudsburg, PA, USA, 2009. Association for Computational Linguistics.

[8] Y. Liu, X. Huang, A. An, and X. Yu. Modeling and predicting the helpfulness of online reviews. In Proceedings of the 2008 Eighth IEEE International Conference on Data Mining, pages 443-452, 2008.