Voice of the Customers: Mining Online Customer Reviews for Product Feature-based Ranking

Kunpeng Zhang Ramanathan Narayanan
Alok Choudhary
Electrical Engineering and Computer Science Department
Northwestern University
2145 Sheridan road, Evanston, IL, 60208, USA

Abstract

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[1] allow customers to add reviews of products they have purchased. These reviews have become a diverse and reliable source to aid other customers. Traditionally, many customers have used expert rankings which rate limited a number of products. Existing automated 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 paper, 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.

[1]http://www.amazon.com

1 Introduction

The rapid proliferation of Internet connectivity has led to increasingly large volumes of electronic commerce. A study conducted by Forrester Research[1] predicted that e-commerce and retail sales in the US market during 2008 were expected to reach $204 billion, an increase of 17% over the previous year. More customers are turning towards online shopping because it is convenient, reliable, and fast. It has become extremely difficult for customers to make their purchasing decisions based only on images and (often biased) product descriptions provided by the seller. Online retailers aim to provide consumers a comprehensive shopping experience by allowing them to choose products based on their specific needs like price, manufacturer, and other attributes. They also allow customers to add reviews of products they have bought. Customer reviews of a product are generally considered more honest, unbiased and comprehensive than descriptions provided by the seller. Furthermore, reviews written by other customers describe the usage experience and perspective of (non-expert) customers with similar needs. A study by comScore and Kelsey group[2] showed that online customer reviews have significant impact on prospective buyers. However, as the number of customer reviews available increases, it is almost impossible for a single user to read all reviews and comprehend them to make informed decisions. Therefore, mining these reviews to extract useful information efficiently is an important and challenging problem.

The abundance of customer reviews available has led to a number of scholars doing valuable and interesting research related to mining and summarizing customer reviews[3, 4, 5, 6, 7, 8]. There has also been considerable work on sentiment analysis of sentences in reviews, as well as sentiment orientation of a review as a whole[9]. There has been relatively little work on ranking products

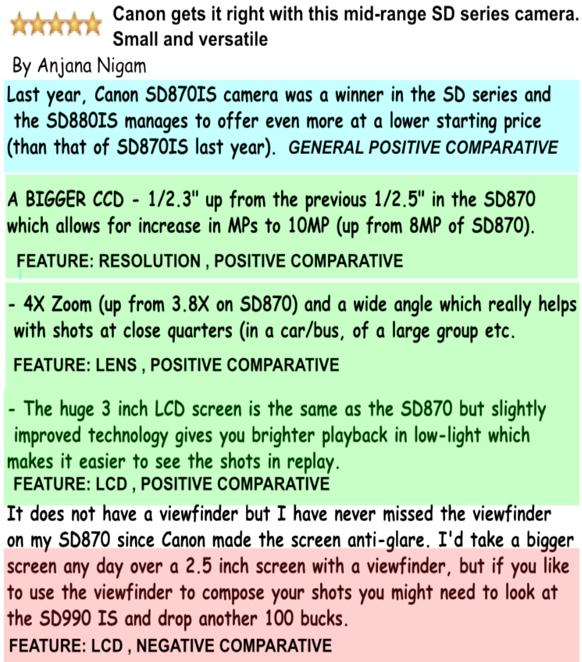

Figure 1: A Customer Review from Amazon.com. Product Features and Subjective/Comparative Sentences are Highlighted.

automatically based on customer reviews. Our 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 less-than-stellar battery life. Thus, both overall ranking and more detailed product feature based ranking are important. In this paper, 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 product's 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. Fig. 1 depicts a typical customer review highlighting the different kinds of sentences. After developing techniques to identify such sentences, we build a weighted and directed product graph(feature-based) that captures the sentiments expressed by customers in reviews. We perform a ranking of products based on the graph.

Particularly, the main contributions in this work are:

- NLP and dynamic programming techniques to identify subjective/comparative sentences(product feature-based) in reviews and determine their sentiment orientations.
- Sentence classification techniques to build a weighted and directed graph which reflects the inherent quality of products in terms of their features.
- A ranking algorithm that uses this massive graph to produce a ranking list of products based on each considered product feature. That is, the end result of the algorithm is a ranking list that a potential customer can use to determine the best products based on customer's importance of one or more product features.

The reminder of this paper is organized as follows. Section 2 contains a summary of related work. We present our techniques in Section 3. Section 4 contains the details of datasets and experimental results followed by the conclusions in Section 5.

2 Related Work

Our work is partly based on and closely related to opinion mining and sentence sentiment classification. Extensive research has been done on sentiment analysis of review text[9, 10] and subjectivity analysis (determining whether a sentences is subjective or objective[22]). Another related area is feature/topic-based sentiment analysis[23], in which opinions on particular attributes of a product are determined. Most of this work concentrate on finding the sentiment associated with a sentence (and in some cases, the entire review). There has also been some research on automatically extracting product features from review text[5]. Though there has been some work in review summarization[3], and assigning summary scores to products based on customer reviews[14], there has been relatively little work on ranking products using customer reviews.

To the best of our knowledge, there has been no focused study on product feature based ranking using customer reviews. The most relevant work is [5], where the authors introduce a unsupervised information extraction system which mines reviews to build a model of important product features, and incorporate reviewer sentiments to measure the relative quality of products. Our work differs from theirs in the following aspects: (1) We use a keyword matching strategy to identify and tag product features in sentences, (2) We use different strategies to assign sentiment orientation to sentences, and incorporate comparisons between different products in our model and (3) Based on sentence classification, we build a product graph for each product feature. By mining this graph, we are able to comprehensively rank products based on reviews they have received.

3 Product Ranking Methodology

3.1 Identifying Product Features

Product features are attributes that provide added functionality to products. In the 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[5]. There has been earlier research on identifying product features and feature sentences, which is not the focus of this paper. We assume that product features associated with a product domain are given. We manually gathered product feature sets for two product cateogories: *Digital Cameras* and *Television* based on the official consumer reports[2]. This is a one-time pre-processing overhead that can be done relatively easily by someone being familiar with a product domain. Table 1 shows the different product features and their major associated synonym sets. 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 described. 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.

[2]http://www.consumerreports.org/

In our review dataset containing 1516001 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, *flash* 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 file 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* means 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/CineMotion* improves the movie-watching experience, which may be important to some users.

3.2 Sentence Labeling

Customers express their opinions about products in multiple ways. We identify two kinds of sentences that are useful while ranking products: Subjective Sentences and Comparative Sentences. Before doing this task, we use MxTerminator[11] to split reviews into sentences because a typical customer review comprises of several sentences. We formally define the different types of sentences below:

Definition 3.1. Subjective Sentence(SS) A sentence expressing direct praise or deprecation about a product. Ex. *This camera has excellent shutter speed.*

Definition 3.2. Comparative Sentence(CS) A sentence which indirectly express an opinion by performing a

Table 1: Keywords Representing 10 Most Important Product Features for *Digital Camera* and *Television* Domains

Digital Camera	TV
resolution|pixel|megapixel	connection|input|output| component video|composite video|HDMI
lens|wide angle|normal range	adjustment|stretch|zoom|expand|compress
optical|zoom|optical zoom|digital zoom	film-mode|frame|theatrical|3:2|pull-down| motion compensation|CineMotion
memory|megabytes|MB	pip|picture-in-picture|dual-tuner|pop| picture-outside-picture|two-tuner
burst|continuous|shutter|recovery|motion|sport	resolution|1080p|1080i|720p
battery|batteries|power	screen|anti-glare|reflectivity|burn-in|shiny| screensaver|pixel-shift
focus|exposure|manual|iso	picture|image|picture quality|image quality
LCD|screen	sound|sound quality|speaker|stereo|audio
compression|compress|jpeg	size|height|width|depth|weight|inch
flash|light	remoter|remote|gear|universal

comparison between two products. Ex. *I think the coolpix is better than the canon sd1200.*

Definition 3.3. Product Comparative Sentence(PCS) A PCS is a comparative sentence and contains at least one product name. Ex. *This TV has much better sound quality when compared to the sony bravia.*

3.3 Identifying Comparative Sentences

There has been some earlier research[16, 17] 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[12], 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 keywords in KW;
- Recognize any words with POS tags $\in$ { JJR(comparative adjective), RBR(comparative adverb), 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[21]. We use a dynamic programming 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 different from the product the sentence is describing while building our ranking model. In [21], we have shown that we get a precision of 82% and a recall of 80% approximately.

3.4 Identifying Sentence Sentiment Orientation

In this section, we describe how we assign sentiment orientations for a sentence. We only consider positive and negative sentiments in this work. Unfortunately, dictionaries and other sources like WordNet[13] do not include sentiment orientation information for each word. Some researchers[18] have used supervised learning algorithms to infer the sentiment orientation of adjectives and adverbs from constraints on conjunctions. [9] contains a summary of existing sentiment analysis techniques. In this paper, we use a simple yet powerful method by utilizing a positive word set(POS) and a negative word set(NEG) developed in the MQPA project[19]. We also add some words of our own. At the end of this process, we get a list of 1974 words for the positive set and 4605 words for the negative set. We use a simple technique to identify the orientation of a sentence using these words. If the sentence contains a word that is in the positive word set, we label this sentence with a positive tag. Negative sentiment words are handled similarly. However, many customers do not like to express their opinions by using assertive sentences but using some negations in their reviews. In this case, the orientation should be switched. We constructed a set of 28 negation words manually. It should be mentioned that we determine the sentence orientation for comparative sentences as well using the same list of sentiment words. A precision of 82% and a recall of 86% approximately can be achieved and shown in [21].

3.5 Constructing the Product Graph

We use the subjective and comparative sentences found to construct a directed and weighted graph for each product feature that can be mined to reveal the relative quality of products. The product graph for feature f is defined as follows: $G_f = (V, E)$ where

- V is a set of nodes, $V = \{p_i \mid$ each node represents a product, $0 < i < n\}$,
- E is a set of node pairs, called arcs or directed edges. An arc $e = (p_i, p_j)$ is considered to be directed from p_i to p_j. $E = \{e_k = (p_i, p_j), \mid W_{e_i}$ is the weight of the edge $e_i, 0 < i, j < n, 0 < k < m\}$,

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

Consider a comparative sentence in the reviews for product P_i describing feature f. If this sentence compares product P_i with product P_j, we add a directed edge from P_j to P_i. The second step is to assign a weight to this edge. A comparative sentence occurring in the reviews for product P_i and comparing it with product P_j is considered a positive comparative($PC(P_i, P_j)$) if it implies that P_i is better than P_j. If it implies that P_i is worse that P_j, it is considered a negative comparative($NC(P_i, P_j)$). For each edge(P_j, P_i), we count the number of positive (PC) and negative (NC) comparative sentences associated with the pair (P_i, P_j) respectively. We assign the ratio PC/NC as the weight of the edge linking P_j to P_i. The weight of a node is used to represent the inherent quality of a product. For a node P_i, we use the ratio of the number of positive(PS) and negative(NS) subjective sentences (PS/NS) as its weight.

3.6 Ranking Products

We evaluate the relative importance of each product according to feature f by using the pRank algorithm described below. The classic PageRank algorithm[20] treats all edges equally and does not take node weights into account. A node has a higher importance if it is pointed to from relatively important nodes. In pRank, we not only consider the relative importance among products, but also take the importance/quality of the product itself into account. This means that the node

Table 2: Product Ranking Results for G_f in Fig. 2

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

weight is also crucial to the ranking, in addition to the edge weights. The idea can be formalized using the equation below:

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

- pRank(P) is the product ranking of product P;
- pRank(P_i) is the product ranking of product P_i and n is the number of incoming links on product P;
- $\mathbf{1}_{\{P_i,P\}}$ is an indicator function, s.t.

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

- $C_e(P_i) = \frac{W_e(P_i,P)}{\sum_{j=1}^{m} W_e(P_i,P_j)}$, 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.

$PS_f(A) = 1, PS_f(B) = 2, PS_f(C) = 3, PS_f(D) = 4$
$NS_f(A) = 3, PC_f(B, A) = 3, PC_f(B, C) = 7$
$PC_f(B, D) = 3, PC_f(A, C) = 2, NC_f(B, C) = 2$

Based on these sentence statistics, we build a product graph G_f(see Fig. 2). Edge weights are determined by comparative sentences, and node weights are determined by subjective sentences. Since the reviews of product C have 7 positive comparative sentences mentioning product B (and feature F), and 2 negative comparative sentences mentioning B (and feature f), there is an edge from C to B with weight 3.5. It must be mentioned that to prevent edges with infinite length (when the number of negative comparative sentences is 0), we set the minimum value of the denominator to 1 while

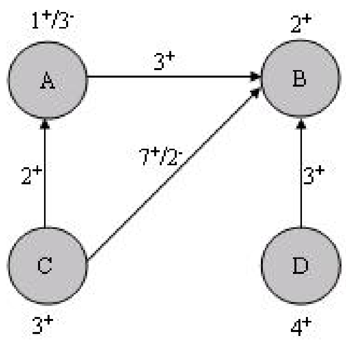

Figure 2: A Simple Ranking Example With a Product Graph G_f Having 4 Products Regarding a Specific Feature f ('+' means positive and '-' means negative)

computing edge weights. By using our algorithm, we get the ranking score for each node shown in the Table 2. The ranking order(the smaller, the product better) for this graph is $B \rightarrow D \rightarrow C \rightarrow A$. 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. Algorithm 1 below summarizes our ranking methodology.

Algorithm 1 pRank Rank Products for Feature f

Require: Product Feature(f), Product Category(Cat).
Ensure: The ranking list of products belonging to category Cat for the feature f .

XML_Data = Download(Cat);
SENT = Get_Sentence(XML_Data);
LSENT = Label(SENT, F);
$\{PS, NS, PC, NC\} \leftarrow LSENT$;
for each sentence $s \in \{PC, NC\}$ **do**
  Find all product comparison pairs $\{p_i, p_j\}$ using dynamic programming;
  Pair $\leftarrow \{p_i, p_j\}$ + $'Pos'\ or\ 'Neg'$;
end for
Count $PS_{p_i}, NS_{p_i}, PC_{p_i,p_j}, NC_{p_i,p_j}$ related to all products;
for each product p_i **do**
  for each product p_j **do**
    if $i == j$ **then**
      $Matrix[i,i] = PS_{p_i}/NS_{p_i}$;
    else
      $Matrix[i,j] = PC_{p_j,p_i}/NC_{p_j,p_i}$;
    end if
  end for
end for
Ranking List = Rank($Matrix[]$);
return Ranking List;

4 Experiment Results

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.com[3] and BrowseNodes.com[4]. The Digital Camera dataset contains 83005 reviews (for 1350 products) and Television dataset contains 24495 reviews (for 760 products) collected by August, 2009. Table 3 and Table 4 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 the product graph. There is no filtering done for product features. We then mine this overall graph $G_{overall}$ using the ranking algorithm described in Section 3.6. 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. More details about this evaluation can be found in[21]. In this paper, 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 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(Ex. lens), then the probability that it ranks in the

[3] http://www.amazon.com/
[4] http://www.browsenodes.com/

Table 3: Breakdown of Subjective/Comparative Sentences(Digital Camera)

Feature/Overall	No. of Sentences	No. of Subjective Sentences		No. of 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

Table 4: Breakdown of Subjective/Comparative Sentences(TV)

Feature/Overall	No. of Sentences	No. of Subjective Sentences		No. of Comparative Sentences	
		Positive	Negative	Positive	Negative
Sound	13877	1599	1933	456	303
Screen	9021	1374	1457	501	344
Size	7214	492	516	342	214
Connection	6299	465	641	239	163
Resolution	6155	286	306	418	256
Picture Quality	4987	2847	1750	201	65
Remoter	4554	619	715	175	117
Adjustment	1704	170	215	74	48
PIP	1205	139	175	49	43
Film-Mode	1022	167	158	53	23
TV	460610	17843	28510	10224	9162

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.

Another aspect of our feature ranking methodology is the relative importance of different product features. A valid question that can be asked is: Which product features are customers looking for when making their choices? To answer this question we define two metrics, Relative Feature Fraction(RFF) and Importance of Feature(IF).

Definition 4.1. Relative Feature Fraction: $RFF_f = \frac{N_f}{\sum_f N_f}$*100%, where N_f is the number of sentences labeled with feature f.

Definition 4.2. Importance of Feature: $IF_f = \frac{|X \cap Y_f|}{|X|} * 100$, where $X = \{$top 10% of overall ranked products$\}$, and $Y_f = \{$top 10% of products according to feature $f\}$.

Table 5 shows the RFF for different features in Digital Camera and Television categories respectively. We observe that customers are mentioning *flash* and *sound* (for Camera and TV) product features in their reviews most often. However, a higher RFF does not always indicate greater importance of a feature. The importance of

Table 5: The Relative Feature Fraction(RFF_f) for Digital Camera and TV

Digital Camera	RFF_f	**TV**	RFF_f
Flash	18.18%	Sound	24.76%
Battery	15.96%	Screen	16.10%
Focus	15.93%	Size	12.87%
Lens	13.67%	Connection	11.24%
Optical	10.77%	Resolution	10.98%
LCD	9.72%	Picture Quality	8.90%
Resolution	5.63%	Remoter	8.13%
Burst	5.40%	Adjustment	3.04%
Memory	4.06%	PIP	2.15%
Compression	0.67%	Film-Mode	1.82%

Table 6: Importance of Feature(IF_f) for Digital Camera and TV

Digital Camera Features	IF_f	**TV** Features	IF_f
Lens	79.9	Size	78.7
Resolution	79.8	Film-Mode	72.3
Optical	77.5	Picture Quality	70.7
Focus	76.3	Connection	69.1
Memory	76.3	PIP	69.1
Burst	75.2	Sound	67.5
Lcd	74.1	Remoter	67.5
Flash	72.9	Adjustment	64.3
Battery	71.6	Screen	61.2
Compression	68.4	Resolution	61.2

Table 7: Top 10 Rated Digital Cameras for Each Feature and Overall Ranking

Lens	Resolution	Optical	Focus
PENTAX K100D SUPER	CANON POWERSHOT SD200	CASIO EXILIM PRO	CASIO EXILIM PRO
KODAK EASYSHARE ONE	NIKON COOLPIX S4	KODAK EASYSHARE ONE	KODAK EASYSHARE ONE
CANON POWERSHOT SD500	NIKON COOLPIX L3	CANON POWERSHOT SD550	PANASONIC DMC-FX37A 10.1MP
CANON POWERSHOT SD990IS	NIKON COOLPIX S600	CANON POWERSHOT SD990IS	CANON POWERSHOT S60
NIKON COOLPIX L3	KODAK EASYSHARE C653	NIKON COOLPIX 3200	CANON POWERSHOT SD990IS
NIKON COOLPIX 8400	NIKON COOLPIX P90	NIKON COOLPIX L3	NIKON COOLPIX 3200
NIKON COOLPIX L1	NIKON COOLPIX L4	NIKON COOLPIX S500	PENTAX OPTIO P70
KODAK EASYSHARE C653	SONY ALPHA A700K	HP PHOTOSMART M537	NIKON COOLPIX L3
NIKON COOLPIX P6000	NIKON COOLPIX P50	NIKON COOLPIX P50	NIKON COOLPIX 8400
NIKON COOLPIX L4	PENTAX OPTIO 60	NIKON COOLPIX 7600	NIKON COOLPIX L1
Memory	**Burst**	**LCD**	**Flash**
KODAK EASYSHARE ONE	CASIO EXILIM PRO	KODAK EASYSHARE ONE	KODAK EASYSHARE ONE
CANON POWERSHOT S80	CANON POWERSHOT SD500	PANASONIC DMC-FX37A 10.1MP	CASIO EXILIM PRO
PANASONIC DMC-FX37A 10.1MP	PANASONIC DMC-FX37A 10.1MP	CANON POWERSHOT S80	PANASONIC DMC-FX37A 10.1MP
PENTAX OPTIO A10	CANON POWERSHOT S80	PENTAX OPTIO A10	NIKON COOLPIX S200
CANON POWERSHOT SD990IS	CANON POWERSHOT S60	CANON POWERSHOT SD990IS	SONY DSCP150 7MP
NIKON COOLPIX 3200	CANON POWERSHOT SD990IS	NIKON COOLPIX 3200	CANON EOS 1D
OLYMPUS SP-550UZ 7.1MP	NIKON COOLPIX 3200	CANON POWERSHOT S400	NIKON COOLPIX 8400
NIKON COOLPIX 4300	NIKON COOLPIX 995	CANON POWERSHOT G3	NIKON COOLPIX L1
CANON POWERSHOT S410	CANON POWERSHOT S100	SONY DSCP150 7MP	NIKON COOLPIX P90
NIKON COOLPIX 8400	NIKON COOLPIX L3	NIKON COOLPIX S500	SONY ALPHA A700K
Battery	**Compression**	**Overall Quality**	
CANON POWERSHOT SD990IS	CANON POWERSHOT A620	NIKON COOLPIX L1	
NIKON COOLPIX S4	CANON POWERSHOT SD300	PANASONIC DMC-FX37A 10.1MP	
NIKON COOLPIX S500	KODAK EASYSHARE ZD710	CANON POWERSHOT A990IS	
NIKON COOLPIX L3	CANON POWERSHOT S100	NIKON COOLPIX P90	
KODAK EASYSHARE C653	NIKON COOLPIX 8700	HP PHOTOSMART M537	
HP PHOTOSMART M537	NIKON COOLPIX 4300	CANON POWERSHOT A70	
NIKON COOLPIX P90	NIKON COOLPIX L3	NIKON COOLPIX 8800	
NIKON COOLPIX P6000	NIKON COOLPIX P50	FUJIFILM FINEPIX A330	
NIKON COOLPIX L4	CANON POWERSHOT S230	KODAK EASYSHARE C653	
SONY ALPHA A700K	CANON POWERSHOT A95	OLYMPIS STYLUS 550	

a feature should reflect the role a feature plays in influencing a customer. The IF metric measures the agreement between the overall ranking and feature-specific ranking. The values for IF for different product features are shown in Table 6. They indicate that our product feature-based ranking is consistent with overall ranking (which is verified previously by comparing with expert ranking [21]). In addition, this ratio should not be close to 1 because overall quality is different from feature quality. Table 6 also tells us that lens (Digital Camera) and size (Television) are the leading factors influencing the overall quality of a product. Table 7 shows the top 10 cameras according to the overall ranking and feature-specific ranking. These results are consistent with the arguments made earlier.

5 Conclusion

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.

6 Acknowledgement

This work was supported in part by NSF grants CNS-0551639, HECURA CCF-0621443, OCI 0956311, SDCI OCI-0724599, OCI-0956311, IIS-0905205, and DOE grant DE-FG02-08ER25848.

References

[1] Forrester research, `http://www.comscore.com/Press_Events/Press_Releases/2007/11/Online_Consumer_Reviews_Impact_Offline_Purchasing_Behavior`.

[2] comscore and Kelsey, `http://www.shop.org/c/journal_articles/view_article_content?groupId=1&articleId702&version=1.0`.

[3] M. Hu and B. Liu, *Mining and Summarizing Customer Reviews*, Proceedings of the 10th ACM International Conference on Knowledge Discovery and Data Mining (SIGKDD-2004), 8 (2004), pp. 168–174.

[4] M. Hu and B. Liu, *Mining Opinion Features in Customer Reviews*, Proceedings of the 19th National Conference on Artificial Intelligence., 7 (2004), pp. 755-760.

[5] A. Popescu and O. Etzioni, *Extracting product features and opinions from reviews*, Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing., (2005), pp. 339-346.

[6] B. Liu, M. Hu, and J. Cheng, *Opinion Observer: Analyzing and Comparing Opinions*, WWW., 5 (2005), pp. 342-351.

[7] S. Kim, P. Pantel, T. Chklovski, and M. Pennacchiotti, *Automatically Assessing Review Helpfulness*, EMNLP., 7 (2006), pp. 423-430.

[8] B. He, C. Macdonald, J. He, and I. Ounis, *An Effective Statistical Approach to Blog Post Opinion Retrieval*, CIKM., 10 (2008), pp. 1063-1069.

[9] B. Liu, Sentiment Analysis and Subjectivity, to appear in Handbook of Natural Language Processing, Second Edition, 2010.

[10] B. Pang, L. Lee, and S. Vaithyanathan, *Thumbs up? Sentiment Classification using Machine Learning Techniques*, EMNLP., (2002), pp. 79-86.

[11] J. Reynar and A. Ratnaparkhi, *A Maximum Entropy Approach to Identifying Sentence Boundaries.*, In Proceedings of the Fifth Conference on Applied Natural Language Processing , March 31-April 3, 1997. Washington, D.C.

[12] X. Phan, *CRFTagger: CRF English POS Tagger*, `http://crftagger.sourceforge.net/`, 2006.

[13] G. A. Mille, C. Fellbaum, etc., *WordNet: An Electronic Lexical Database*, 5 (1998),MIT Press.

[14] M. Hu, and B. Liu, *Opinion extraction and summarization on the web*, AAAI., 7 (2006), pp. 1621-1624.

[15] B. Liu, X. Li, W. S. Lee, and P. S.Yu, *Text Classification by Labeling Words*, AAAI., 2004, pp. 425-430.

[16] N. Jindal, and B. Liu, *Identifying Comparative Sentences in Text Documents*, SIGIR., 8 (2006), pp. 244-251.

[17] N. Jindal, and B. Liu, *Mining Comparative Sentences and Relations*, AAAI., (2006), pp. 1331-1336.

[18] V. Hatzivassiloglou, and K. R. McKeown, *Predicting the semantic orientation of adjectives*, Proceedings of the eighth conference on European chapter of the Association for Computational Linguistics, 1997, pp. 174-181.

[19] MPQA corpus `http://www.cs.pitt.edu/mpqa`, 2002.

[20] L. Page, S. Brin, R. Motwani, and T. Winograd, *The PageRank citation ranking: Bringing order to the Web*, Stanford University, 1999.

[21] K. Zhang, R. Narayanan, and A. Choudhary, *Mining Online Customer Reviews for Ranking Products*, Technical Report, EECS department, Northwestern University, 2009.

[22] S. Arora, M. Joshi, and C. P. Rose, *Identifying Types of Claims in Online Customer Reviews*, NAACL HLT., 6 (2009), pp. 37-40.

[23] Y. Lu, C. Zhai, and N. Sundaresan, *Rated Aspect Summarization of Short Comments*, WWW., 4 (2009), pp. 131-140.