CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

ORCA – Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository:https://orca.cardiff.ac.uk/id/eprint/68390/

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Jin, Jian, Ji, Ping and Liu, Ying 2014. Prioritising engineering characteristics based on customer online reviews for quality function deployment. Journal of Engineering Design 25 (7-9) , pp. 303-324. 10.1080/09544828.2014.984665

Publishers page: http://dx.doi.org/10.1080/09544828.2014.984665

Please note:
Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See http://orca.cf.ac.uk/policies.html for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.

Prioritizing Engineering Characteristics based on Customer Online Reviews for QFD

Journal:	*Journal of Engineering Design*
Manuscript ID:	CJEN-2014-0219.R2
Manuscript Type:	Article
Keywords:	quality function deployment < Design Methods, Techniques and Tools, Knowledge and Information Management, User review analysis, Customer reviews, QFD

For Peer Review Only

SCHOLARONE™
Manuscripts

Prioritizing Engineering Characteristics based on Customer Online Reviews for QFD

Jian Jin[a], **Ping Ji**[b] and **Ying Liu**[c*]

[a] Department of Information Management,
School of Government,
Beijing Normal University,
Beijing 100875, China

[b] Department of Industrial and Systems Engineering,
The Hong Kong Polytechnic University, Hong Kong SAR, CHINA

[c*] Institute of Mechanical and Manufacturing Engineering,
School of Engineering,
Cardiff University, Cardiff CF24 3AA, United Kingdom
LiuY81@cardiff.ac.uk

For Peer Review Only

ABSTRACT:
In market-driven product design, customer requirements are usually obtained from consumer surveys. However, valuable customer requirements can also be found in a large number of online reviews. Largely due to their free text nature and the quantity, these reviews are often neglected and are seldom utilized directly by designers. In this work, one important question in quality function deployment (QFD) on how to prioritize engineering characteristics is investigated. Customer opinions concerning engineering characteristics are extracted from online reviews. By taking advantage of such opinion information, an ordinal classification approach is proposed to prioritize engineering characteristics. In a pairwise manner, in which customer opinions are deemed as features and the overall customer satisfactions are regarded as the target values, the weights of engineering characteristics are derived. Furthermore, an integer linear programming model is implemented to convert the pairwise results into the original customer satisfaction ratings. Finally, an exploratory case study is presented using reviews of four branded printers collected from Amazon and their analysis was conducted by two experienced design engineers. The experimental study reveals the merits of the proposed approach.

Keywords: Customer reviews; user review analysis; QFD, engineering characteristics; product design.

1. INTRODUCTION

Online reviews contain valuable information about customer opinions and customer requirements (CRs). This information benefits both potential consumers and product designers (Liu, Lu and Loh, 2007). It is highly possible that potential consumers are affected by online reviews before they make decisions and product designers will learn valuable customer needs from online reviews. However, for some hot products, there may be hundreds of or even thousands of customer online reviews and they are distributed in different websites, which makes it difficult to handle all relevant reviews. This problem has been widely discussed in the research area of computer science. Some researchers developed innovative models to extract the opinions of consumers from online reviews (Hu and Liu, 2004; Liu, Hu and Cheng, 2005). Generally, it is named as sentiment analysis or opinion mining. Their research focus is mainly on sentiment identification (Hassan and Radev, 2010; Zagibalov and Carroll, 2008; Lin and He, 2009), sentiment extraction (Qiu et al., 2009; Kim and Hovy, 2005), and opinion retrieval (Zhang et al., 2008; Lin et al., 2010) at different levels. However, none of these research studies discuss how online reviews benefit product designers in terms of customer analysis. In the design area, only a few studies concern about the value of online reviews, such as how to summarize and analyze the topic structure of online reviews for product design (Zhan, Loh and Liu, 2009), how to identify helpful online reviews from the perspective of product designers (Liu et al., 2013), etc. What they neglect is how to assemble their findings into product design directly. For

For Peer Review Only

instance, how to improve the current product models according to the requirements in online reviews.

For example, in Figure 1, one typical iPad review is illustrated.. It complains that "…I was hoping the third gen wouldn't be noticeably thicker and heavier than the 2, but unfortunately it was. I could definitely tell the difference when reading ebooks, which I do a lot…". As seen from this figure, this consumer is not satisfied with the size and the weight of iPad, which leads to an overall three-star is given to express his/her overall customer satisfaction. This review provides suggestive information to designers. Accordingly, when launching the next generation, designers may consider, with limited time and budget, which parts they need to improve in order to satisfy potential consumers. More specially, in the process of new product design, designers need to prioritize several engineering characteristics (ECs), such as the size, the weight, etc. It is also an indispensable step in Quality Function Development (QFD).

For Peer Review Only

1,651 of 1,738 people found the following review helpful

★★★☆☆ **Good, but ...**, March 25, 2012

By **Bookenz**

This review is from: **Apple iPad MD328LL/A (16GB, Wi-Fi, White) 3rd Generation (Personal Computers)**

I've been a big iPad fan and was waiting anxiously for the third generation to come out. I bought the original iPad, and when the 2 came out, I happily sold the 1 and upgraded. I was thrilled with the thinner, lighter and improved iPad 2. So naturally, when word came out that the third one was about to be released in March 2012, I was right on board to buy one. Sold my 2 and preordered the 3 from Apple the day it was released.

I was hoping the third gen wouldn't be noticibly thicker and heavier than the 2, but unfortunately it was. I could definitely tell the difference when reading ebooks, which I do a lot. I couldn't really tell any difference between the speed and clarity of the 3, but to be fair, I didn't compare the two models side by side. I've no doubt the 3 is superior in this regard. I don't use the cameras, so don't care about this since I have a very nice digital camera for that.

The one thing about the third generation iPad that concerned me was the heat issue. Shortly after receiving it, I was reading an ebook and noticed the left side was warm. Not hot, but definitely warm enough for me to notice it. This reminded me of laptops I've had that have overheated and shut down, and here I was only reading a book. My iPad 2 never had this problem and I used it a lot. I also had some difficulty backing up the 3 to the cloud, again, something that wasn't ever a problem with the 2.

After reading some reviews of others experiencing the heating problem with the latest iPad, and really missing the thinner and lighter iPad 2 , I decided to return the iPad 3 to my local Apple store and buy a new iPad 2. They had no problem taking it back and I was glad to see the 2 had come down in price. When the sales associate asked me why I was returning the 3, I told him about the heat problem. He didn't seem surprised and said it was because the 3 has a larger battery.

The third gen has a faster CPU and retina dsplay, but I never thought the 2 had any problems with speed, and the clarity of the display has always seemed fine to me. But I'm certainly no expert. I'm very happy I made the exchange and enjoying using my iPad 2 again. For my uses--books, internet, email and games, the cooler, thinner and lighter iPad 2 is a better choice for me.

By the way, I just want to add that I think it's unfair to compare the Kindle Fire and Nook to the iPad. These devices are just ebook readers and they're fine for that--I've tried them both. Rather like saying, okay, you can buy this Mercedes or you can get the Volkswagen. Once you've used an iPad, you're totally spoiled and trying to browse the Internet with a Kindle or Nook is an exercise in frustration after using an iPad.

Update January 7, 2014 - since writing this review for the iPad 3, which I ended up returning, I've moved on to the iPad Air, which is noticeably lighter and faster than the iPad 2.

Help other customers find the most helpful reviews | Report abuse | Permalink

Was this review helpful to you? Yes No

Figure 1. One typical online review

Correspondingly, in this research, the focus is on how to prioritize ECs based on the CRs in online reviews. Notice that, in the field of engineering design, the term, "engineering characteristics", is "also known as technical attributes, product technical requirements or design requirements" and they "describe the product in the language of the engineer" (Sener and Karsak, 2011). Also, ECs is named as "the voice of the company" and they are employed to "determine how well the company satisfies the needs." For such reasons, they are often established at an early design stage according to the company's strategic goals (Kahraman, Ertay and Buyukozkan, 2006). Thus,

generally, ECs can be design specifications, substitute quality characteristics, company's strategic goals, engineering attributes, government regulations, or specification practice standards (Wang, 1999; Chan and Wu, 2005). Specifically, in this research, the term refers to the translation of CRs to an engineering language from the perspective of product designers, which can be generally deemed as how product designers deem and evaluate these CRs.

Different from many conventional methods, in this research, online reviews, rather than survey data, are taken into consideration to identify CRs. Specially, customer opinions about ECs are firstly extracted from online reviews. Then, the overall ratings affiliated with online reviews are taken as the overall customer satisfaction about products. Combined with the overall customer satisfaction, these opinions in online reviews are utilized to prioritize ECs. In particular, an ordinal classification approach is proposed based on pairwise algorithms about learning to rank. The idea about the marginal maximization of pairwise algorithms is borrowed in this ordinal classification approach, which intends to derive the weights of ECs. Moreover, in order to transform the pairwise results to the original ratings of customer satisfaction, an integer linear programming model is formulated, which has not been explained clearly by many pairwise algorithms. In the same time, a running example is presented to explain the proposed ordinal classification approach and the integer linear programming model step by step. Finally, a case study of printer design is prepared to show some interesting findings and how the proposed method assist designers to decide critical ECs.

For Peer Review Only

The contributions of this research are at least three folds. Firstly, online reviews, which are fundamentally different from customer survey data, are utilized to prioritize ECs. It is the first attempt to make online reviews directly utilized by product designers. Secondly, the overall customer satisfaction and opinions over ECs in online reviews are important for the prioritization of ECs. Accordingly, an ordinal classification approach is proposed to prioritize ECs by the analysis of customer reviews. An integer linear programming model is also formulated to transform the pairwise results from the ordinal classification approach to the original customer satisfaction ratings. Finally, comparative case studies are conducted to expose interesting phenomena. Frequently talked ECs might not necessarily become a deciding factor for the new product design, while some details of products may influence the overall customer satisfaction.

The rest of this paper is organized as follows: Section 2 introduces the related work. In Section 3, the problem to be investigated in this research is formally defined. Section 4 describes the research efforts and illustrates the proposed technical approaches. Section 5 presents the details of a case study and discusses some experimental results. Section 6 concludes this research.

2. LITERATURE REVIEW

2.1. Importance Weighting of Engineering Characteristics

Due to the time and budget limitations, when designers conceive to improve product models, it is usually unreasonable to consider ECs without any bias. Importance

For Peer Review Only

weighting of ECs, which is one important problem in QFD, becomes crucial for the resource allocation as well as the final decision-making in new product design. A nonlinear programming model was proposed to prioritize ECs in fuzzy environments (Wang and Chin, 2011). Two numerical examples were shown to verify the availability of this model. A fuzzy weighted average method was also proposed to prioritize ECs in fuzzy QFD (Chen, Fung, and Tang, 2006). In this method, a discrete solution was obtained by changing the fuzzy weighted average problem to a pair of fractional programming problem for each EC. Kwong et al. (2011) argued that both the human perception and the customer heterogeneity were found to influence the importance of ECs in QFD, but most of the relevant researches only center at one of them. Accordingly, a fuzzy group decision-making method combining both a fuzzy weighted average method and an ordinal ranking was proposed to incorporate the two influential factors. This approach was argued to be better than the method proposed by Chen, Fung and Tang (2006).

Some researchers argued that the importance of ECs in QFD can be evaluated from two aspects, namely, the needs of customer aspect and the needs of manufacturer aspect (Geng et al., 2010). From the perspective of customer needs, the analytic network process was utilized to estimate the initial importance of ECs by considering the relationships of customer needs, product characteristics, and service related ECs. The fuzzy set theory was then applied in the analytic network process to deal with the uncertainty in decision-making. From the perspective of manufacturer needs, the data envelopment analysis was employed to adjust the initial weights of

For Peer Review Only

product characteristics by considering both the business competition and the implementation difficulty. QFD was regarded as a grey system in (Li, Zhang and Gao, 2009). The relationships between CRs and ECs were determined by the grey relational matrix. A grey method was then utilized to prioritize ECs. Moreover, Kano's Model was also seen to be integrated with QFD to recognize the importance of ECs (Chaudha, et al., 2011).

Some research studies associating with customer satisfaction and ECs are also valuable to be highlighted, although these do not handle the problem about importance weighting of ECs straightforwardly. For instance, a Neuro fuzzy approach was also reported to generate a customer satisfaction model (Kwong, Wong and Chan, 2009). An example for notebook computer design is given to show that this model is better than a statistical regression approach. A genetic programming model is proposed to associate ECs with customer satisfaction, in which the interactions between ECs as well as high-order relationship between ECs and customer satisfaction are considered (Chan, Kwong and Wong, 2011). Then, an orthogonal least-squares algorithm is employed to decide the coefficients of the relationship. A fuzzy least-squares regression method is also proposed to model the fuzzy relationship between customer satisfaction and ECs as well as the fuzzy relationship among ECs (Kwong et al., 2010) .

Notice that, although many approaches were developed regarding prioritizing ECs, in the most of publications, survey data often become the source of CRs. However, online reviews are fundamentally different from survey data and their

For Peer Review Only

differences are at least in three folds. First, due to the time and budget limitations, it is usually time-consuming and labor-intensive for designers to gain a large number of survey data manually. But a large number of customer reviews are always available online. Second, survey data are usually generated from questionnaires with specific intentions and some customers are invited to select the most proper answers from a given list. But customer reviews are free text. Some of them may contain only a few words, while, in the other case, hundreds, or even thousands, of words may appear in one review. Sufficient customer concerns can be found from online reviews. Thirdly, survey data seldom contain personal opinions, while online reviews are one important way to express their opinions. These differences make many approaches which are proposed on the basis of survey data in the design area cannot be utilized directly to handle CRs in a large volume of online reviews.

2.2. Online Reviews for Product Design

Although online reviews are widely accepted as one important source of CRs, only a few studies discuss the value of online reviews in product design.

A text mining system, where online reviews were integrated with the domain knowledge, was reported on knowledge discovery and management in product design (Liu, Lu and Loh, 2007). To aggregate CRs from online reviews for product design, a framework was presented (Decker and Trusov, 2010). This framework was utilized to infer the relative effect of product features and the effect of different brands on the overall customer satisfaction. An automatic summarization approach was seen to

For Peer Review Only

analyze the topic structure of online reviews (Zhan, Loh and Liu, 2009). This approach was utilized to discover and assemble important topics in online reviews. The final summary of multiple reviews was then clustered by the topic structure and different clusters were ranked according to the importance of different topics.

Online reviews were also reported to be utilized in the prediction about product design trends (Tucker and Kim, 2011). Sentiment polarity of product features were extracted from online reviews. Then, the Holt-Winters exponential smoothing method was employed to model product preference trends. A system that monitor customer opinions from textual data is built (Goorha and Ungar, 2010). First, frequent phrases and phrases near the terms of interest are extracted from textual data. These phrases are then utilized to identify which of them appear dramatically. Also, whether a phrase is regarded as an interesting one is depended on how often they are referred, how often they are referred comparing with before and how specific they refer to a topic. These results are illustrated by an interactive user interface. Also, in order to present the results effectively, TFIDF weights and the cosine similarity method is utilized to cluster relevant terms.

A three steps method is proposed for customer driven product design selection by the analysis of online reviews (Wang et al., 2011). In the first step, product attributes were extracted from online reviews. In the second step, a hierarchical customer preference model was built by Bayesian linear regression. Product ratings, category ratings, attribute ratings and product specifications are taken into consideration in this hierarchical model. Finally, an optimization problem was

For Peer Review Only

formulated to maximize potential profit by taking engineering constraints into consideration. In (Liu et al., 2013), the helpfulness of online reviews was initially defined in the viewpoint of designers for QFD. According to designers' arguments, four categories of features are extracted from online reviews. With these features, the helpfulness of online reviews is inferred by a regression method. Also, categories of experiments confirm that, without domain-dependent features, there is no significant loss in terms of helpfulness prediction for online reviews from the perspective of product designers.

2.3. Pairwise Approaches of Learning to Rank

Learning to rank is a type of supervised learning task. The objective is to construct a ranking model from training data, which sorts new objects according to their degrees of relevance (Joachims et al., 2007). The algorithms of learning to rank are generally classified into three types, pointwise approaches, pairwise approaches and listwise approaches. Different input and output spaces, hypothesis and loss functions are defined in different types. In this research, a pairwise approach is proposed to prioritize ECs based on customer reviews. To understand this approach, some relevant pairwise algorithms are reviewed in this subsection.

The input space of a pairwise approach contains a pair of documents, which is represented as feature vectors. For the output space, a pairwise approach contains the pairwise preference between each pair of documents. The hypothesis space of a pairwise approach contains functions, which take a pair of documents as input and

For Peer Review Only

output the relative order between them. The loss function of pairwise approaches measures the inconsistency between the predicted relationship and the ground truth labels of document pairs.

Pairwise approaches do not target at accurately predicting the relevance degree of each single document, and only care about the relative order between two documents. The learning procedure of pairwise approaches is conducted over document pairs. An example of outputs may take values from positive one and negative one, which indicates the pairwise preference between each pair of documents. A conceptual illustration for pairwise approaches of Learning to Rank is proposed in Tie-Yan Liu's tutorial in *SIGIR'08*, which is shown in Figure 2 (Liu, 2008).

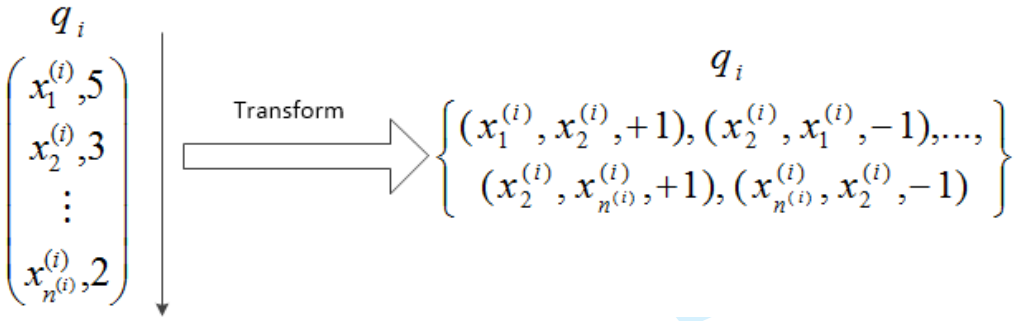


Figure 2. An conceptual illustration for pairwise approaches of learning to rank (Liu, 2008)

In Figure 2, a typical training set for learning to rank is shown, which consists some training queries, their associate documents as well as the corresponding relevance judgments. For each training query q_i, the associate documents are represented by feature vectors $x^{(i)} = \{x_j^{(i)}\}_{j=1}^{n^{(i)}}$, where $n^{(i)}$ is the number of documents associated with query q_i. However, one transformation needs to be conducted to build a training set for the pairwise approach. The input space of the pairwise approach are

represented by feature vectors $\{x_j^{(i)}, x_k^{(i)}\}_{j \neq k}^{n^{(i)}}$, which contains all pairs of documents, and the output space is the value of positive one and negative one, which indicate pairwise preference between each pair of documents. Take two documents ($x_2^{(i)}$, 3) and ($x_{n^{(i)}}^{(i)}$, 2) for instance. In the pairwise approach, according to the corresponding relevance judgments 5 and 3, they are transformed to ($x_2^{(i)}, x_{n^{(i)}}^{(i)}$,1) and ($x_{n^{(i)}}^{(i)}, x_2^{(i)}$,-1)

Accordingly, in pairwise approaches, ranking is usually reduced to a classification problem on document pairs. It makes many conventional algorithms for classification applicable for this problem. A neural network was built to learn a preference function for all possible document pairs in training data (Burges et al., 2005; Rigutini et al., 2011). A boosting approach over document pairs was utilized to combine ranking functions in RankBoost (Freund et al., 2003). Based on SVM, RankSVM was proposed to perform the pairwise classification (Herbrich, Graepel and Obermayer, 2000; Joachims, 2002). RankSVM differs from SVM at its constraint part and the loss function, which was built from document pairs. However, one hyperplane is employed by RankSVM, which is argued to be hard to handle complex ranking problems (Qin et al., 2007). Multiple hyperplanes were proposed to train a ranking model for document pairs. Finally, the ranking results predicted by each ranking model were aggregated to the final ranking result.

2.4. A Short Summary

In the design field, various models were developed based on survey data.. But online reviews which contain valuable information about CRs are not concerned. The

For Peer Review Only

differences between online reviews and survey data make many conventional methods in product design not be applicable to handle CRs in a big volume of distributed and fast-evolving online reviews. Although a few studies discuss how to utilize online reviews in product design, as one of critical steps in QFD, the importance weighting of ECs for QFD based on online customer reviews is neglected.

A large number of customer reviews which represent consumer opinions and requirements are observed online. Although positive online reviews do not always bring new consumers, negative reviews may lead to the loss of potential customers. Accordingly, it becomes critical to conduct the analysis of CRs efficiently and make an immediate and effective response, especially in a fierce competitive market. As one downside of online reviews, not all of these data contain much valuable information to product designers. In the previous research (Liu et al., 2013), the helpfulness of online reviews is perceived, evaluated and predicted from the perspective of product designers. These techniques in the previous research enable product designers to concentrate only on high-quality customer reviews. Hence, in this research, based on high-quality customer online reviews, one of important questions about how to prioritize ECs for QFD is centered. It will be highly possible to facilitate designers' work in the process of conceptual design.

3. PROBLEM DEFINITION

In QFD, taking customer needs as input, product designers usually need to prioritize ECs. To clarify the problem in this research, some notation will be defined.

For Peer Review Only

In some e-commerce websites, such as Amazon.com, each product has a series of high quality customer reviews, r_1, r_2, ..., r_p. These reviews contain valuable customer needs. However, due to the time limitation and the budget cost, on the basis of customer needs, it is unreasonable to consider all of ECs without bias. Typically, a list of ECs includes $<ec_1, ec_2 ... ec_n>$.

Typically, customers may present different opinions over different ECs. For example, in a particular review r_i, a customer is satisfied or unsatisfied with ec_j. It can be denoted as (ec_j, O_{ij}). O_{ij} is the associated opinion about ec_j in r_i. Accordingly, r_i can be represented as an opinion vector $O_i = <O_{i1}, O_{i2}, ..., O_{in}>$. A positive O_{ij} denotes that, in r_i, this consumer is satisfied with ec_j, while, a negative value denotes that an unsatisfied opinion. Also, in r_i, there may be zero, one or more sentences associated with a specific ec_j. If a consumer does not mention ec_j in r_i, the corresponding O_{ij} is assumed to be zero, which implies that this consumer has a neutral opinion. For the case that there are more than one sentences discussing ec_j in r_i, the average value of O_{ij} is taken as the consumer's final opinion on ec_j.

In addition, consumers are also encouraged to give a rating towards the overall satisfaction. For example, in Figure 1, a three-star iPad review is shown. It is denoted as, in review r_i, a rating cs_i is utilized to express the overall satisfaction. According to the definition of cs_i and O_i, r_i can be (O_i, cs_i). For a customer review set, containing p reviews, it can be denoted as $< (O_1, cs_1), (O_2, cs_2) ... (O_p, cs_p)>$. Now, based on the information about customer opinions and the overall satisfaction, the central question is how to prioritize ECs for QFD.

For Peer Review Only

Specially, it is about how to infer the weights $W = <w_1, w_2, ..., w_n>$ of ec_1, ec_2, ..., ec_n from $<(O_1, cs_1), (O_2, cs_2), ..., (O_p, cs_p)>$ by exploiting all p reviews. Mathematically, it can be denoted as:

$$cs_i = f(\sum_{j=1}^{n} w_j O_{ij}) \quad (1)$$

$f(x)$ is a devised function transforming the sum of weighted opinions over ECs into the overall customer satisfaction. Notice that, the objective here is not to find or settle the exact form of such function but to denote the weights of opinions over different ECs are to be learned according to opinions in customer online reviews.

It appears that a regression model is qualified to learn w_1, w_2... w_n. But it is arguable since regression models are generally utilized to analyze questions with continuous values as the target, while, in this research, cs_i is a discrete value. Thus, the classification model might be more persuasive than regression models. It is still questionable since the inherent ranking information of cs_i will be neglected by simple classification models. For example, in one review, a five-star is given by a consumer. Suppose that it is predicted as a four-star by model one and a three-star by model two. In this scenario, model one is favorable since it is closer to the original five-star, though two models lead to the same prediction error with simple classification models. Another potential technique is learning to rank. However, models of learning to rank neglect that objects can be possibly placed in the same position. For example, two reviews may present the same customer satisfaction rating on products while learning to rank fail to consider such case. Hence, models of learning to rank are also not suitable to be utilized directly to prioritize ECs based on customer reviews.

For Peer Review Only

Particularly, in this research, both the classification and the ranking information should be taken into consideration, when considering how to prioritize ECs based on customer reviews. Hence, an ordinal classification model is required.

4. PRIORITIZING ENGINEERING CHARACTERISTICS BASED ON CUSTOMER REVIEWS

4.1 An Ordinal Classification Approach

As presented in Section 2, one famous algorithm of learning to rank is RankSVM, which is a pairwise approach. In RankSVM, the weights of features W are learned from training data, which enables the distance between hyperplanes of document pairs is maximized. W is then utilized to predict the preference of document pairs. In this research, if opinions over different ECs are regarded as features and the overall customer satisfaction is regarded as the expected ranking position, the weights will be learned accordingly. The limitations is that, RankSVM stresses two documents are well separated. It is not true for the customer satisfaction ratings of different reviews because they may be given the same level of customer satisfaction. However, the idea about learning W from document pairs and maximizing the distances of hyperplanes is highly instructive. Accordingly, in this research, learning the weights W of ECs from online reviews is transformed to learn W from review pairs. Hence, at first, a set of review pairs P should be derived from the whole review set D. P should contain all set of review pairs $((O_i, cs_i), (O_j, cs_j))$.

For Peer Review Only

4.1.1. An Example of Nine Reviews with Six Engineering Characteristics

To clarify the proposed ordinal classification approach clearly, a running example is introduced through this section. In this example, there are nine printer reviews, r_1, r_2, ..., r_9. One to five stars are utilized to indicate the overall customer satisfaction about the printer. It is denoted as $cs_i \in \{1, 2, 3, 4, 5\}$. The values of customer satisfaction for all nine reviews are supposed in Table 1.

Table 1. The values of customer satisfaction for nine reviews

cs_1	cs_2	cs_3	cs_4	cs_5	cs_6	cs_7	cs_8	cs_9
5	2	3	4	5	2	4	1	3

With these nine reviews, six important ECs, ec_1, ec_2, ..., ec_6, are concerned only. For instance, they may be "ease of use", "noise", "print quality", "Wifi integration", "duplex printing" and "card slot". The corresponding weights are denoted as w_1, w_2, ..., w_6, which are to be learned from the nine customer reviews. But notice that other ECs may be considered in this example. The objective to select some exemplary ECs is to illustrate how the weights are learned by the proposed ordinal classification approach, not to take these exact six ECs only into considerations. Moreover, opinions over ec_1, ec_2, ..., ec_6, may be different. In this research, a five-degree metric, $O_{ij} \in \{-2, -1, 0, 1, 2\}$, is employed to evaluate the opinion, where "–2" stands for the least satisfied and "2" for the most satisfied. Accordingly, O_i $=<O_{i1}, O_{i2}, ..., O_{i6}\}>$, will be the customer satisfaction over ec_1, ec_2, ..., ec_6 in a review r_i. These opinions are planned randomly as illustrated in Table 2.

Table 2. The opinions over six ECs of nine reviews

	r_1	r_2	r_3	r_4	r_5	r_6	r_7	r_8	r_9

For Peer Review Only

ec_1	1	-2	2	-2	-1	1	1	-2	-1
ec_2	0	-1	1	1	1	0	-1	0	-1
ec_3	1	-2	-2	1	0	-2	1	-2	-2
ec_4	2	2	0	1	-2	-1	-1	-2	-1
ec_5	-1	2	-2	2	1	-1	-2	0	0
ec_6	-1	1	-1	-2	-2	-1	1	-2	2
cs_i	5	2	3	4	5	2	4	1	3

4.1.2. Building the review pair set P

As mentioned, in this research, the weights of ECs will be learned from review pairs. Hence, a review pair set P should be deduced accordingly.

Firstly, for review r_i and r_j, if $cs_i > cs_j$, or equivalently, r_i is ranked better than r_j, $(O_i - O_j, 1)$ is put into P. If $cs_i < cs_j$, then $(O_i - O_j, -1)$ is put into P. If $cs_i = cs_j$, it means r_i is ranked equivalently to r_j, then $(O_i - O_j, 0)$ is put into P. Then, in the nine reviews example, there are $\binom{9}{2} = \frac{9 \times 8}{2} = 36$ review pairs. All review pairs in P are shown in Table 3.

Table 3. The set P for the nine reviews

r_1, r_2, (3 1 3 0 -3 -2, 1)	r_1, r_3, (-1 -1 3 2 1 0, 1)	r_1, r_4, (3 -1 0 1 -3 1, 1)
r_1, r_5, (2 -1 1 4 -2 1, 0)	r_1, r_6, (0 0 3 3 0 0, 1)	r_1, r_7, (0 1 0 3 1 -2, 1)
r_1, r_8, (3 0 3 4 -1 1, 1)	r_1, r_9, (2 1 3 3 -1 -3, 1)	r_2, r_3, (-4 -2 0 2 4 2, -1)
r_2, r_4, (0 -2 -3 1 0 3, -1)	r_2, r_5, (-1 -2 -2 4 1 3, -1)	r_2, r_6, (-3 -1 0 3 3 2, 0)

For Peer Review Only

r_2, r_7, (-3 0 -3 3 4 0, -1)	r_2, r_8, (0 -1 0 4 2 3, 1)	r_2, r_9, (-1 0 0 3 2 -1, -1)
r_3, r_4, (4 0 -3 -1 -4 1, -1)	r_3, r_5, (3 0 -2 2 -3 1, -1)	r_3, r_6, (1 1 0 1 -1 0, 1)
r_3, r_7, (1 2 -3 1 0 -2, -1)	r_3, r_8, (4 1 0 2 -2 1, 1)	r_3, r_9, (3 2 0 1 -2 -3, 0)
r_4, r_5, (-1 0 1 3 1 0, -1)	r_4, r_6, (-3 1 3 2 3 -1, 1)	r_4, r_7, (-3 2 0 2 4 -3, 0)
r_4, r_8, (0 1 3 3 2 0, 1)	r_4, r_9, (-1 2 3 2 2 -4, 1)	r_5, r_6, (-2 1 2 -1 2 -1, 1)
r_5, r_7, (-2 2 -1 -1 3 -3, 1)	r_5, r_8, (1 1 2 0 1 0, 1)	r_5, r_9, (0 2 2 -1 1 -4, 1)
r_6, r_7, (0 1 -3 0 1 -2, -1)	r_6, r_8, (3 0 0 1 -1 1, 1)	r_6, r_9, (2 1 0 0 -1 -3, -1)
r_7, r_8, (3 -1 3 1 -2 3, 1)	r_7, r_9, (2 0 3 0 -2 -1, 1)	r_8, r_9, (-1 1 0 -1 0 -4, -1)

Take r_1 and r_2 for example. As seen from Table 2, cs_1 equals to 5 and cs_2 equals to 2. cs_1 is bigger than cs_2, then ($O_1 - O_2$, 1) is put into P. $O_1 - O_2$ is calculated as $O_1 - O_2$ = (1 - (-2), 0 - (-1), 1 - (-2), 2 - 2, (-1) - 2, (-1) - 1). For r_2 and r_3, since cs_2 is smaller than cs_3, then ($O_2 - O_3$, –1) is put into P. Similarly, it is calculated as ((-2)-2, (-1)-1, (-2) - (-2), 2-0, 2-(-2), 1-(-1)). The third example is between r_1 and r_5. Since both cs_1 and cs_5 equal to 5, ($O_1 - O_5$, 0) is put into P.

With P defined, deriving the weights of ECs based on customer reviews turns to be a tri-classification problem, which attempts to classify review pairs to {–1, 0, 1}. Let one review pair is (OP_k, cr_k), where OP_k equals to $O_i - O_j$. OP_k can be also denoted as $<OP_{k1}, OP_{k2}, ..., OP_{kn}>$, where $OP_{ks} = O_{is} - O_{js}$. cr_k is a discrete value, which represents the relationship of cs_i and cs_j. Taking OP_k as the feature vector and cr_k as the target class, it is a tri-classification problem to derive the weights of ECs.

However, if a further step is taken, a binary classification will be shown. Notice that, when $cs_i < cs_j$, if ($O_j - O_i$, 1) is put into P, rather than ($O_i - O_j$, –1), a

For Peer Review Only

binary classification is shown. Clearly, it does not affect the value of W. But the problem is that the ranking relationship of two reviews is lost. Thus, to trace which review presents a higher level of customer satisfaction, whether $cs_i > cs_j$ or $cs_j > cs_i$ is kept in practice. Now OP_k is either $O_i - O_j$ or $O_j - O_i$, and cr_k is "1" or "0". "1" denotes that two reviews receive different levels of customer satisfaction, while "0" denotes two reviews receive the same. Taking the previous nine reviews as examples, accordingly, all review pairs in P become what are shown in Table 4.

Table 4. Review pairs in P after applying transformation rules

r_1, r_2, (3 1 3 0 -3 -2, 1)	r_1, r_3, (-1 -1 3 2 1 0, 1)	r_1, r_4, (3 -1 0 1 -3 1, 1)
r_1, r_5, (2 -1 1 4 -2 1, 0)	r_1, r_6, (0 0 3 3 0 0, 1)	r_1, r_7, (0 1 0 3 1 -2, 1)
r_1, r_8, (3 0 3 4 -1 1, 1)	r_1, r_9, (2 1 3 3 -1 -3, 1)	r_3, r_2, (4 2 0 -2 -4 -2, 1)
r_4, r_2, (0 2 3 -1 0 -3, 1)	r_5, r_2, (1 2 2 -4 -1 -3, 1)	r_2, r_6, (-3 -1 0 3 3 2, 0)
r_7, r_2, (3 0 3 -3 -4 0, 1)	r_2, r_8, (0 -1 0 4 2 3, 1)	r_9, r_2, (1 0 0 -3 -2 1, 1)
r_4, r_3, (-4 0 3 1 4 -1, 1)	r_5, r_3, (-3 0 2 -2 3 -1, 1)	r_3, r_6, (1 1 0 1 -1 0, 1)
r_7, r_3, (-1 -2 3 -1 0 2, 1)	r_3, r_8, (4 1 0 2 -2 1, 1)	r_3, r_9, (3 2 0 1 -2 -3, 0)
r_5, r_4, (1 0 -1 -3 -1 0, 1)	r_4, r_6, (-3 1 3 2 3 -1, 1)	r_4, r_7, (-3 2 0 2 4 -3, 0)
r_4, r_8, (0 1 3 3 2 0, 1)	r_4, r_9, (-1 2 3 2 2 -4, 1)	r_5, r_6, (-2 1 2 -1 2 -1, 1)
r_5, r_7, (-2 2 -1 -1 3 -3, 1)	r_5, r_8, (1 1 2 0 1 0, 1)	r_5, r_9, (0 2 2 -1 1 -4, 1)
r_7, r_6, (0 -1 3 0 -1 2, 1)	r_6, r_8, (3 0 0 1 -1 1, 1)	r_9, r_6, (-2 -1 0 0 1 3, 1)
r_7, r_8, (3 -1 3 1 -2 3, 1)	r_7, r_9, (2 0 3 0 -2 -1, 1)	r_9, r_8, (1 -1 0 1 0 4, 1)

Notice that, in SVM or RankSVM, "–1" and "1" are often utilized to denote two different classes and the parallel hyperplanes. Accordingly, rather than "0", "–1"

For Peer Review Only

is deployed to symbolize two reviews with the same level of customer satisfaction. cr_k is then either "–1" or "1". The transformation rules from the whole review set D to review pair set P is summarized as:

$$\begin{aligned} (O_i - O_j, 1) &\rightarrow \text{P}, \quad \text{if } cs_i > cs_j \\ (O_j - O_i, 1) &\rightarrow \text{P}, \quad \text{if } cs_i < cs_j \\ (O_i - O_j, -1) &\rightarrow \text{P}, \quad \text{if } cs_i = cs_j \end{aligned}$$

With the transformation rules, the set P that are derived from the nine review example is shown in Table 5.

Table 5. The final review pairs in P

r_1, r_2, (3 1 3 0 -3 -2, 1)	r_1, r_3, (-1 -1 3 2 1 0, 1)	r_1, r_4, (3 -1 0 1 -3 1, 1)
r_1, r_5, (2 -1 1 4 -2 1, -1)	r_1, r_6, (0 0 3 3 0 0, 1)	r_1, r_7, (0 1 0 3 1 -2, 1)
r_1, r_8, (3 0 3 4 -1 1, 1)	r_1, r_9, (2 1 3 3 -1 -3, 1)	r_3, r_2, (4 2 0 -2 -4 -2, 1)
r_4, r_2, (0 2 3 -1 0 -3, 1)	r_5, r_2, (1 2 2 -4 -1 -3, 1)	r_2, r_6, (-3 -1 0 3 3 2, -1)
r_7, r_2, (3 0 3 -3 -4 0, 1)	r_2, r_8, (0 -1 0 4 2 3, 1)	r_9, r_2, (1 0 0 -3 -2 1, 1)
r_4, r_3, (-4 0 3 1 4 -1, 1)	r_5, r_3, (-3 0 2 -2 3 -1, 1)	r_3, r_6, (1 1 0 1 -1 0, 1)
r_7, r_3, (-1 -2 3 -1 0 2, 1)	r_3, r_8, (4 1 0 2 -2 1, 1)	r_3, r_9, (3 2 0 1 -2 -3, -1)
r_5, r_4, (1 0 -1 -3 -1 0, 1)	r_4, r_6, (-3 1 3 2 3 -1, 1)	r_4, r_7, (-3 2 0 2 4 -3, -1)
r_4, r_8, (0 1 3 3 2 0, 1)	r_4, r_9, (-1 2 3 2 2 -4, 1)	r_5, r_6, (-2 1 2 -1 2 -1, 1)
r_5, r_7, (-2 2 -1 -1 3 -3, 1)	r_5, r_8, (1 1 2 0 1 0, 1)	r_5, r_9, (0 2 2 -1 1 -4, 1)
r_7, r_6, (0 -1 3 0 -1 2, 1)	r_6, r_8, (3 0 0 1 -1 1, 1)	r_9, r_6, (-2 -1 0 0 1 3, 1)
r_7, r_8, (3 -1 3 1 -2 3, 1)	r_7, r_9, (2 0 3 0 -2 -1, 1)	r_9, r_8, (1 -1 0 1 0 4, 1)

4.1.3. Learning the Weights of Engineering Characteristics

For Peer Review Only

Now, the question turns to learn the weight W to classify review pairs in P into two classes ("–1" and "1"). It is similar to what has been done in SVM. In SVM, W denotes the weights of features. It should be tuned to maximize the distance between the parallel hyperplanes and there are no additional constraints about W. But, in this research, W is defined as the weights of ECs and it should be nonnegative. Also, in SVM, if two classes cannot be separated, a compromised idea is employed. A slack variable, ξ_k, is chosen to measure the degree of compromise. This idea is adopted to devise this ordinal classification method. Accordingly, the model is shown as follows:

$$\begin{aligned} &\min \sum \xi_k + \frac{C}{2} W^T W \\ &s.t.\ \hat{cr}_k = \sum_{j=1}^{n} w_j OP_{kj} \\ &\hat{cr}_k cr_k \geq 1 - \xi_k \\ &W \geq \mathbf{0} \\ &\xi_k \geq 0 \end{aligned} \tag{2}$$

Model (2) presents the details about the ordinal classification approach. It is adapted from SVM, where an additional constraint of weights W is considered. ξ_k is the slack variable, which is utilized to estimate the degree of compromise. $\frac{1}{2} W^T W$ is the regularization term, which is to control the over-fitting phenomenon. The coefficient C governs the relative importance of the regularization term compared with the sum of compromise terms. The linear term, $\sum_{j=1}^{n} w_j OP_{kj}$, is to estimate the customer satisfaction relationship between two reviews $\hat{cr}_k$. The distance between two hyperplanes is $2 - 2\xi$. It intends to be made as large as possible, which makes two classes to be discriminated. In the third constraint, the weight w_i of ec_i is restrained to be bigger than or equal to zero. Bold "**0**" denotes a zero vector, rather than a scalar

For Peer Review Only

zero. Accordingly, the weights over six ECs, w_1, w_2, ..., w_6, of the nine reviews are calculated as:

$$\begin{aligned} &\min \sum_{k=1}^{36} \xi_k + \frac{C}{2} \sum_{t=1}^{6} W_t^2 \\ s.t.\ &\hat{cr}_k = \sum_{j=1}^{6} W_j OP_{kj} \\ &\hat{cr}_k cr_k \geq 1 - \xi_k \\ &W \geq \mathbf{0} \\ &\xi_k \geq 0 \end{aligned} \tag{3}$$

According to the data presented in Table 5 and the minimization problem of Model (3), W will be equal to W = (0.1154, 1.5000, 0.8077, 0.0000, 0.1923, 0.8462) if C is set to 0.5. Moreover, the predicted customer satisfaction relationship $\hat{cr}_k$ of all pairs is illustrated in Table 6.

Table 6. The inferred customer satisfaction of 36 review pairs

r_1, r_2, 1	r_1, r_3, 1	**r_1, r_4, -1**	**r_1, r_5, 1**	r_1, r_6, 1	r_1, r_7, 1
r_1, r_8, 1	r_1, r_9, 1	r_3, r_2, 1	r_4, r_2, 1	r_5, r_2, 1	**r_2, r_6, 1**
r_7, r_2, 1	r_2, r_8, 1	r_9, r_2, 1	r_4, r_3, 1	r_5, r_3, 1	r_3, r_6, 1
r_7, r_3, 1	r_3, r_8, 1	**r_3, r_9, 1**	**r_5, r_4, -1**	r_4, r_6, 1	**r_4, r_7, 1**
r_4, r_8, 1	r_4, r_9, -1	r_5, r_6, 1	**r_5, r_7, -1**	r_5, r_8, 1	r_5, r_9, 1
r_7, r_6, 1	r_6, r_8, 1	r_9, r_6, 1	r_7, r_8, 1	r_7, r_9, 1	r_9, r_8, 1

Compared with the ground truth data presented in Table 5, there exist seven pairs are not correctly inferred. They are highlighted with bold characters.

4.2 Transforming the Results to the Original Customer Satisfaction Rating

For Peer Review Only

In the previous section, the weights of ECs are learned by the pairwise classification approach. However, a new interesting question will come out if the pairwise approach is applied.

Suppose the proposed pairwise approach is expected to be evaluated by some classification metrics like *Precision*, *Recall* as well as *F-measure* and some ranking metrics like Mean Average Precision (*MAP*) and Normalized Discounted Cumulative Gain (*NDCG*). Hence, the predicted level of customer satisfaction of each review is required. But, in a pairwise approach, only the relationship between two instances is obtained. The exact position, or the level of customer satisfaction of individual review, in this research, is still unknown. If the proposed ordinal classification is perfect, the level of customer satisfaction can be gained from the pairwise results. However, it is generally hard to train a classifier without any errors, and there are possibly some misclassified instances. As seen in the previous nine review example, there are some misclassified review pairs by using the pairwise approach, so it is impossible to transform the pairwise results into the levels of customer satisfaction faithfully. For example, there are three reviews, r_i, r_j, and r_k. The customer satisfaction relationship are predicted by Model (2) as $\hat{cs}_i > \hat{cs}_j$, $\hat{cs}_j > \hat{cs}_k$, and $\hat{cs}_k > \hat{cs}_i$. Conflicting results lead to that the levels of customer satisfaction of these three reviews are not able to be assigned to satisfy all the predicted relationship. Hence, the interesting question is how to transform the pairwise results to the original level of customer satisfaction, in which the predicted relationship is satisfied, or, in which the number of violations about the predicted relationship is minimized.

For Peer Review Only

In particular, two reviews r_i and r_j are predicted as $\hat{cs}_i > \hat{cs}_j$. The question is how to assign the level of customer satisfaction to meet the predicted relationship. If a Boolean variable α denotes whether the relationship is satisfied or not, either zero or one, the problem can be formulated mathematically by Model (4):

$$\begin{aligned} &\hat{cs}_i - \hat{cs}_j \geq 1 - M \cdot \alpha \\ &\hat{cs}_j - \hat{cs}_i \geq 1 - M \cdot (1-\alpha) \\ &\alpha \in \{0,1\} \\ &\hat{cs}_i, \hat{cs}_j \in \{1,2,3,4,5\} \end{aligned} \tag{4}$$

In Model (4), if the predicted relationship $\hat{cs}_i > \hat{cs}_j$ is satisfied, α equals to zero, otherwise α equals to one. M is a large number, for instance M equals to 10^3. Likewise, if $\hat{cs}_i < \hat{cs}_j$, the equivalent model is as follows:

$$\begin{aligned} &\hat{cs}_j - \hat{cs}_i \geq 1 - M \cdot \beta \\ &\hat{cs}_i - \hat{cs}_j \geq 1 - M \cdot (1-\beta) \\ &\beta \in \{0,1\} \\ &\hat{cs}_i, \hat{cs}_j \in \{1,2,3,4,5\} \end{aligned} \tag{5}$$

If $\hat{cs}_i = \hat{cs}_j$, γ is symbolized whether the equation relationship is satisfied or not. The model is:

$$\begin{aligned} &\hat{cs}_i - \hat{cs}_j \leq M \cdot \gamma \\ &\hat{cs}_j - \hat{cs}_i \leq M \cdot \gamma \\ &\gamma \in \{0,1\} \\ &\hat{cs}_i, \hat{cs}_j \in \{1,2,3,4,5\} \end{aligned} \tag{6}$$

According to Model (4) ~ (6), α, β, and γ denote whether the corresponding relationship is satisfied or not. Hence, the sum of α, β, and γ represents the total number of the unsatisfied relationship.

For Peer Review Only

Now, the question of transforming the pairwise results into the customer satisfaction rating faithfully turns to minimize the sum of α, β, and γ. Combining Model (4) ~ (6), the Model to derive $\hat{cs}_i, \hat{cs}_j, \hat{cs}_k$ is:

$$\begin{aligned} & \min\{\sum \alpha_i + \sum \beta_j + \sum \gamma_k\} \\ s.t. \quad & \hat{cs}_{ai} - \hat{cs}_{bi} \geq 1 - M \cdot \alpha_i \\ & \hat{cs}_{bi} - \hat{cs}_{ai} \geq 1 - M \cdot (1 - \alpha_i) \\ & \hat{cs}_{bj} - \hat{cs}_{aj} \geq 1 - M \cdot \beta_j \\ & \hat{cs}_{aj} - \hat{cs}_{bj} \geq 1 - M \cdot (1 - \beta_j) \\ & \hat{cs}_{ak} - \hat{cs}_{bk} \leq M \cdot \gamma_k \\ & \hat{cs}_{bk} - \hat{cs}_{ak} \leq M \cdot \gamma_k \\ & \alpha_i, \beta_j, \gamma_k \in \{0,1\} \\ & \hat{cs}_{ai}, \hat{cs}_{bi}, \hat{cs}_{aj}, \hat{cs}_{bj}, \hat{cs}_{ak}, \hat{cs}_{bk} \in \{1,2,3,4,5\} \end{aligned} \tag{7}$$

As seen from Model (7), it is an integer linear programming optimization problem and it is solvable to obtain optimal results.

For the previous example of nine reviews, as noted in Table 6, there are 13 pairs of $\hat{cs}_i > \hat{cs}_j$, 19 pairs of $\hat{cs}_j > \hat{cs}_i$, and 4 pairs of $\hat{cs}_i = \hat{cs}_j$. Then, Model (7) will be utilized to obtain the levels of customer satisfaction of nine reviews from the relationship of 36 review pairs. The details are illustrated in Model (8).

$$\begin{aligned} & \min\{\sum_{i=1}^{13} \alpha_i + \sum_{j=1}^{19} \beta_j + \sum_{k=1}^{4} \gamma_k\} \\ s.t. \quad & \hat{cs}_{ai} - \hat{cs}_{bi} \geq 1 - M \cdot \alpha_i \quad \forall i \in [1,13] \\ & \hat{cs}_{bi} - \hat{cs}_{ai} \geq 1 - M \cdot (1 - \alpha_i) \\ & \hat{cs}_{bj} - \hat{cs}_{aj} \geq 1 - M \cdot \beta_j \quad \forall j \in [1,19] \\ & \hat{cs}_{aj} - \hat{cs}_{bj} \geq 1 - M \cdot (1 - \beta_j) \\ & \hat{cs}_{ak} - \hat{cs}_{bk} \leq M \cdot \gamma_k \quad \forall k \in [1,4] \\ & \hat{cs}_{bk} - \hat{cs}_{ak} \leq M \cdot \gamma_k \\ & \alpha_i, \beta_j, \gamma_k \in \{0,1\} \\ & \hat{cs}_{ai}, \hat{cs}_{bi}, \hat{cs}_{aj}, \hat{cs}_{bj}, \hat{cs}_{ak}, \hat{cs}_{bk} \in \{1,2,3,4,5\} \end{aligned} \tag{8}$$

For Peer Review Only

According to Model (8), the predicted levels of customer satisfaction of nine reviews are $\hat{cs}_1 = 4, \hat{cs}_2 = 2, \hat{cs}_3 = 3, \hat{cs}_4 = 5, \hat{cs}_5 = 4, \hat{cs}_6 = 2, \hat{cs}_7 = 5, \hat{cs}_8 = 1, \hat{cs}_9 = 3.$

In Model (2) and Model (7), how to prioritize ECs based on customer online reviews and how to transform the pairwise results to the original customer satisfaction ratings are introduced. Overall, the procedure of two phases is shown in Figure 3.

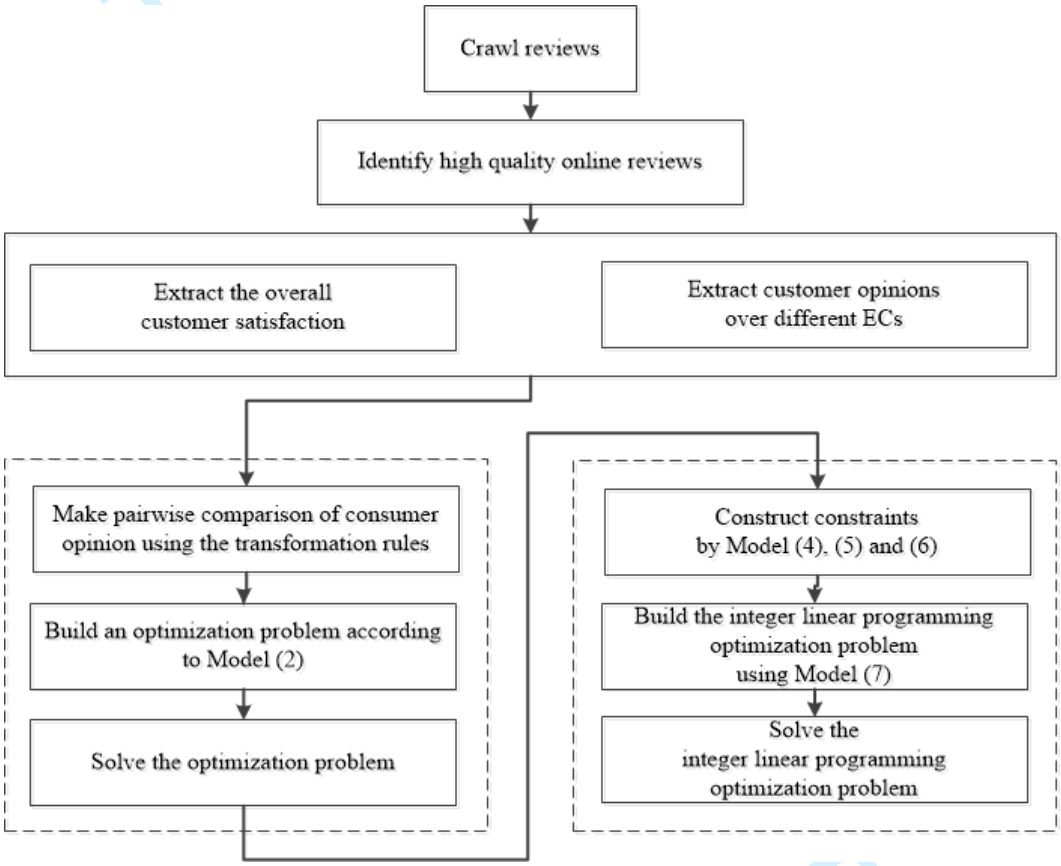


Figure 3. the procedure of prioritizing ECs from online reviews

5. EXPERIMENTAL STUDY AND DISCUSSIONS

5.1. An Exploratory Case Study

To understand how online reviews are evaluated by product designers, an exploratory case study was conducted. In this case study, two customer service clerks were hired

to evaluate online reviews since it is generally difficult to gain the evaluation from experienced designers for specific products. Two clerks were working in Epson Hong Kong and HP Hong Kong. They are very familiar with printers and they had a sound understanding about customer needs. In their daily work, they often need to report various forms of consumer concerns to requirement analysts who help designers to understand CRs and improve their products. Also, two clerks had experiences with printer design using QFD, which contributes to build a high quality dataset.

A web crawler was then utilized to collect online product reviews from Amazon.com and Epson.com. Indeed, the quality of online reviews is one critical problem before these data are utilized by product designers. In our previous work (Liu et al, 2013), the helpfulness of online reviews is perceived, evaluated and predicted from the perspective of product designers. Moreover, an evaluation model is proposed and what makes online reviews are helpful are analyzed in (Liu et al, 2013). Accordingly, with the proposed techniques in the previous work, in this case study, 770 helpful reviews of four color printers (Epson Artisan 810, Epson WorkForce 610, HP 6500 and HP C309) were selected. For short, “A810”, “W610”, “H6500”, and “C309” are employed. In Table 7, the number of reviews of four printers is shown.

Table 7. Number of reviews

Printer	A810	W610	H6500	C309
Number of reviews	258	169	210	133

At first, a list of ECs was collectively suggested by the two annotators, which is illustrated in Table 8, according to their communications with requirement analysts and the working experience in the dedicated field. As clarified in Section 1, ECs in

For Peer Review Only

this list can be deemed as the description or translation of CRs in online reviews to an engineering language from the perspective of product designers.

Table 8. Engineering characteristics

Printer Housing	Power Supply	Fax Setting	Brand
Wifi Integration	Ease of Setup	Ease of Use	Noise
Duplex Printing	Print Quality	Print Head	Package
Software Updated	Scan Software	LCD Panel	Outlooks
Auto Document Feeder	Printing Speed	Hopper Unit	Card Slot
Supplementary Software	Mac Compatible	Ink Longevity	Durability

Reviews were then labeled by the two customer service clerks. They read all reviews and distinguish which ECs are mentioned. Five discrete values from "–2" to "2" are utilized to denote the customer opinions over ECs. "–2" means the least satisfied and "2" means the most satisfied. An example of one Epson Artisan810 review labeling is shown in Figure 4.

For Peer Review Only

Sentence		Auto Document Feeder	Card Slot	Consumable Replacement	Durability	Ease of setup	Ease of use	Fax Setting	Hopper Unit	Mac Compatible	Ink Longevity	LCD Panel	Noise	Power Supply	Print Quality	Printer Housing	Wifi Integration
I only have the printer for a few days, but so far I am very pleased.	-																
I was a bit nervous before buying because I had read a lot of complai	-																
But the paper handling on my 810 has been flawless, if a tiny bit noi	-												-2				
But the paper handling on my 810 has been flawless, if a tiny bit noi	-								-1								
The actual printing is quiet, and of great quality.	-												1				
The actual printing is quiet, and of great quality.	-														2		
The paper tray feels a bit flimsy, but is easy to remove or insert, a	-								-2								
It can expand to hold legal size paper, and has a separate area for s	-																
The package comes with 3 sheets of Epson high quality 4x6 glossy phot	-														2		
I have been using an HP Wireless printer up until now(model 5850) and	-														2		
Also, they promise that their inks are acid-free and will last 3 or 4	-										0						
Another pre-purchase worry was that there were complaints about the A	-																
Too early to tell, but I 've printed those 3 photos and a fourth on p	-										1						
So, so far so good on ink usage.	-										0						
I was surprised and pleased too that the 810 printer comes with an TW	-			1													
Nice!	-																
Additionally, the inks provided are the same capacity as the refills,	-			2													
A nice feature is that when you plug in a camera memory card in the f	-		2														
Very neat.	-																
It warns on the box that this feature may not work with Mac computers	-									1							
Have n 't tried printing from the iPhone as they claim you can do, tho	-																
I love the big touch screen operation too, it is pretty slick.	-											2					
Made a few copies of some simple pencil drawings and they were very g	-														-1		
Updated Review 03\/2010Here 's a 6 month update.	-																
The printer is getting a lot of work ; we print a lot of classroom ma	-				2												

Figure 4. Review labeling

In this example, the seventh line in Figure 4 is that "the paper tray feels a bit flimsy, but is easy to remove or insert, and there's no fuss to loading your paper in it." This consumer actually complained about the "Hopper Unit", so the annotators wrote "–2" in the corresponding column of the seventh line. If more than one ECs are mentioned in a sentence, the sentence is pasted into the other line and the second item is labeled in a new line. For example, in Figure 4, the fourth sentence is "the actual printing is quiet, and of great quality." "Noise" and "Print Quality", are mentioned. This sentence is labeled repeatedly in order to unambiguously identify the two ECs.

Finally, the customer service clerks double checked these reviews in order to avoid any mislabeling. Given four printer review datasets with manually annotated

Review Only

information as the experimental data, the task is to how to prioritize ECs.

5.2. Results and Discussions

The prioritization of ECs might be regarded to be highly related with the frequency that they are referred to in customer reviews and those frequently mentioned characteristics might be given higher weights. The top five frequently-mentioned ECs are listed in Table 9.

Table 9. Top five frequent engineering characteristics

A810		W610	
Feature	Frequency	Feature	Frequency
Print Quality	0.727	Print Quality	0.488
Ease of Setup	0.495	Wifi Integration	0.377
Scan Software	0.485	Ease of Setup	0.364
Wifi Integration	0.475	Printing Speed	0.364
Hopper Unit	0.364	Noise	0.284
H6500		C309	
Ease of Setup	0.636	Print Quality	0.698
Wifi Integration	0.568	Ease of Setup	0.586
Print Quality	0.426	Wifi Integration	0.483
Scan Software	0.278	Printing Speed	0.302
Noise	0.272	LCD Panel	0.293

As seen from this table, more than 40% consumers prefer talking about "Print Quality", "Ease of Setup" and "Wifi Integration". It is easy to understand. For a printer, the print quality and the usability may always be the first concern. However, whether they should be given a higher priority is unknown. This hypothesis will be examined in the following experiments.

The performance of Model (2) is illustrated in Figure 5. C is the regularization term that avoids the parameters w_i in Model (2) being tuned too large. As seen from Figure 5, the accuracy slopes down gradually with a higher C and, except for the

For Peer Review Only

"W610" dataset, the predicted accuracy is all higher than 70%.

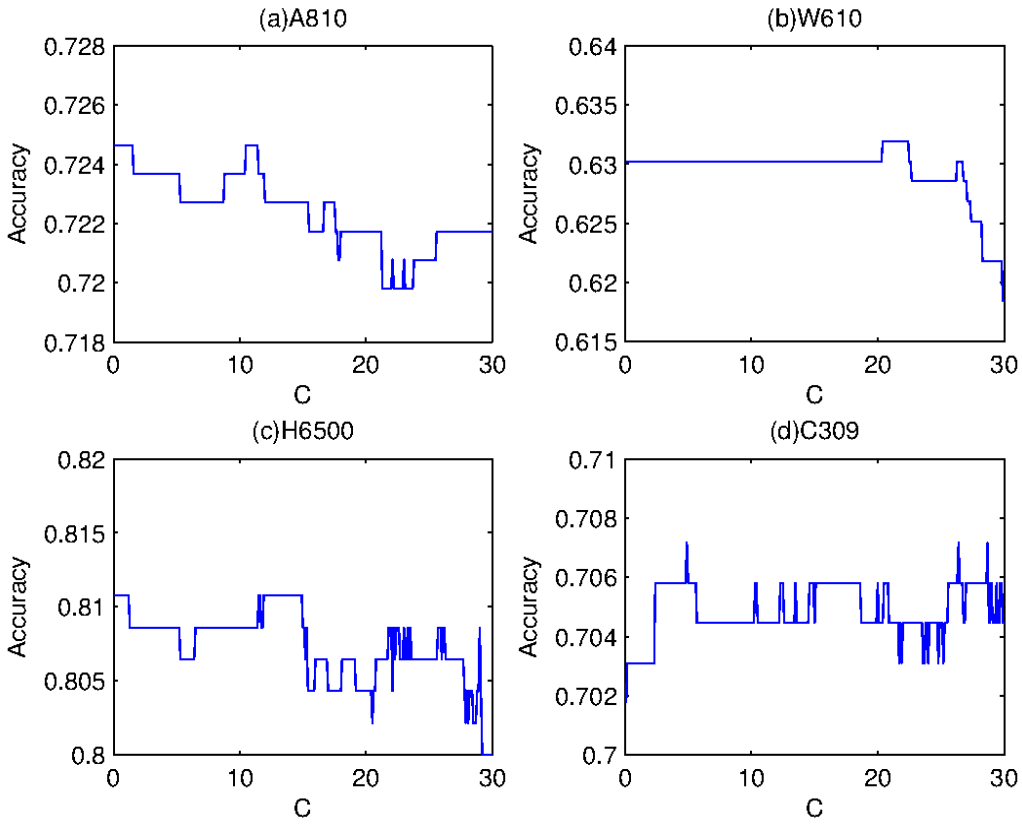


Figure 5. Accuracy vs. regularization term *C*

Notice that the weight might be tuned too large if there is a large proportion of zeros in ECs. In this research, a proportion of zeros illustrates that consumers do not express their opinions or only leave a neutral opinion about the ECs. However, it possibly induces that a higher weight is given to these ECs with zero value in Model (2). It is unreasonable to suggest product designers to make more efforts on those characteristics with little comments. To avoid this problem, in all of the following experiments, those ECs which are mentioned by less than 10% product reviews are neglected.

View Only

According to Model (2), in Table 10, important ECs of the four printer datasets are listed. In all these experiments, *C* equals to 15, where the accuracy is relatively stable in Figure 5. Compared with top frequent ECs in Table 9, somewhat different results are presented.

Table 10. Top important engineering characteristics

	A810	W610
1st	Ink Longevity	Fax Setting
2nd	Mac Compatible	Printing Speed
3rd	Wifi Integration	Wifi Integration
4th	Printing Speed	Ease of Use
5th	Ease of Use	Ease of Setup
	H6500	C309
1st	Ease of Setup	Ease of Setup
2nd	Ease of Use	Wifi Integration
3rd	Wifi Integration	LCD Panel
4th	Scan Software	Noise
5th	Printing Speed	Printing Speed

Firstly, important ECs are suggested in Table 10, when designers plan to improve the current model. Compared with Table 9, frequently mentioned ECs in customer reviews are not necessarily regarded as important ones. For example, according to Table 9, "Print Quality" is frequently discussed, but it does not appear in Table 10 in all of these four printer datasets. It implies that, generally speaking, "Print Quality" is a hot topic in printer reviews, but a higher level of customer satisfaction over "Print Quality" perhaps not necessarily lead to the same level of customer satisfaction about the whole product. It means that product designers are not suggested to pay more attentions on it when they conceive to improve the current model. But it does not mean "Print Quality" is not important for printer design. Although "Print Quality" receives relative lower priority, generally, the high print

For Peer Review Only

quality is considered as a must when a printer is designed. It actually points to another relevant question, how to classify ECs into different categories, such as, must-be, one dimensional and attractive attributes in Kano's Model. This is one future work of this research.

Secondly, important ECs may not be talked about by a large proportion of consumers. Take the "Ease of Use" as an example. This term appears three times in Table 10, but it does not appear in Table 9. It implies that, although this term does not frequently mentioned by consumers, the overall customer satisfaction is impacted by this EC in a certain degree. Product designers are suggested to pay more attention to improve the usability of printers, and the high degree of customer satisfaction may depend on these details.

Thirdly, there are also some common ECs in both Table 10 and Table 9. For instance, "Wifi Integration" and "Ease of Setup" appear in the two tables. These two items, especially "Wifi Integration", are new characteristics for a printer. Without the "Wifi Integration", a printer still works very well. Similarly, with a little complex setting up for some amateur, a printer may still be a good product. However, the user experience will be improved with these creative ECs. These characteristics are the focus of many customer reviews and the overall customer satisfaction is affected by these novel ECs.

Another objective in this research is to explore what are the most important ECs for those consumers who gave a five-star to the product, and whether these ECs are aligned with the ones in Table 10. Thus, some similar experiments were

For Peer Review Only

conducted towards concerning five-star reviews only. Top important ECs are presented in Table 11.

Table 11. Top important engineering characteristics from five star reviews

	A810	W610
1st	Ink Longevity	Fax Setting
2nd	Auto Document Feeder	Wifi Integration
3rd	Consumable Replacement	Printing Speed
4th	Wifi Integration	Ease of Setup
5th	Ease of Use	Noise
	H6500	C309
1st	Mac Compatible	Ease of Setup
2nd	Ease of Use	Supplementary Software
3rd	Wifi Integration	Ink Longevity
4th	Printing Speed	LCD Panel
5th	Print Quality	Print Quality

As seen from Table 11, except for "Wifi Integration", more ECs are not frequently discussed by consumers. The reasons for consumers to give a five-star rating are diversified, compared with the results shown in Table 10. For example, "Ink Longevity" is regarded as an important EC in "A810" and "C309" dataset. But it does not appear in either Table 10 or Table 9. It implies that, although "Ink Longevity" is not frequently mentioned by many consumers and, generally, this EC is regarded as an unimportant one, it should not be neglected if the product is expected to be deemed as a five-star one.

Also, as seen from Table 11, products present different advantages towards how to satisfy consumers to give a five-star rating. For instance, the "Fax Setting" in "W610" dataset is a decisive factor for consumers to make a five-star decision and, similarly, "Mac Compatible" is regarded as the most important EC in the "H6500"

For Peer Review Only

dataset. These phenomena further confirm that it is some details of the product that influence the overall customer satisfaction.

5.3. Performance Evaluation

5.3.1 Evaluation metrics

In Section 4, an ordinal classification approach was proposed to prioritize ECs according to CRs in online reviews, and an integer linear programming problem is formulated to transform the pairwise results into the original customer satisfaction rating. In order to verify the performance of these methods, both classification-based and rank-based performance metrics are examined.

With the integer linear programming model, pairwise results of the proposed ordinal classification approach are transformed to the original level of customer satisfaction. The objective of the integer linear programming model is to enable proposed ordinal classification approach to be testified by standard classification metrics and ranking metrics. *Precision*, *Recall* and F_1 score are widely utilized to evaluate the performance in terms of classification. *Precision* is the fraction of retrieved instances that are relevant, which is calculated as the number of retrieved and relevant results divided by the number of all retrieved results.

$$Precision = \frac{\left|\{relevant\ documents\} \cap \{retrieved\ documents\}\right|}{\left|\{retrieved\ documents\}\right|} \tag{9}$$

The AND operator in the numerator illustrates the intersection of the retrieved document subset and the relevant document subset. *Recall* is the fraction of relevant instances that are retrieved, which is calculated as the number of retrieved and

For Peer Review Only

relevant results divided by the number of all relevant results.

$$Recall = \frac{\left|\{relevant\ documents\} \cap \{retrieved\ documents\}\right|}{\left|\{relevant\ documents\}\right|} \tag{10}$$

F_1 score is the harmonic mean of *Precision* and *Recall*. F_1 score reaches its best value at one and its worst score at zero.

$$F_1 = 2 \cdot \frac{|\ Precision \cdot Recall\ |}{|\ Precision + Recall\ |} \tag{11}$$

To evaluate the proposed ordinal classification approach in terms of ranking metrics, *MAP* and *NDCG* are employed. *MAP* and *NDCG* are often utilized to evaluate the performance in terms of ranking. *MAP* is defined according to the *P@n* metric and *AP*(q). *P@n* shows the precision achieved by considering the top n examples in the ranked list. If there are r_n relevant documents in the top n examples, then $P@n = \frac{r_n}{n}$. *AP*(q) averages the *P@n* over possible values of n. Let r_q be the total number of relevant examples of query q, and $|Q|$ be the total number of examples in query q, and $r(n)$ be a function returning one if the n^{th} ranked example is relevant, and zero, otherwise.

$$\begin{aligned} AP(q) &= \frac{1}{r_q} \sum_{n=1}^{|Q|} P@n \cdot r(n) \\ MAP &= \frac{1}{|\ Q\ |} \sum_{q \in Q} AP(q) \end{aligned} \tag{12}$$

NDCG is to evaluate ranking when multiple levels of relevance are presented. In *NDCG*, the information gain of a document is measured based on its position in the result list. The information gain is accumulated from the top of the result list to the bottom with the gain of each result discounted at lower ranks.

For Peer Review Only

$$DCG = \sum_{i=1}^{n} \frac{\log_2(1+i)}{2^{rel_i} - 1}$$
$$NDCG = \frac{DCG}{IDCG} \quad (13)$$

rel_i is the graded relevance of the result at position i, such as the number of "star" which stands for the level of customer satisfaction in this research. *IDCG* is the normalization term of *DCG*, which ensures that the perfect *NDGC* score for the given set of examples is one. In a faultless ranking algorithm, the *NDCG* will be one.

5.3.2 Experimental evaluations

The performance of the formulated integer linear programming model, which intends to transform the pairwise results into the original customer satisfaction rating with Model (7), is illustrated in Table 12. The final results are evaluated in terms of both classification and ranking metrics. A relative high performance is achieved by the ordinal classification algorithm in all of the four data sets. It proves the availability of the proposed ordinal classification approach.

Table 12. Performance of the ordinal classification algorithm

	Classification			Ranking	
	Precision	*Recall*	F_1	*MAP*	*NDCG*
A810	0.674	0.717	0.695	0.913	0.988
W610	0.600	0.629	0.614	0.889	0.975
H6500	0.710	0.742	0.725	0.953	0.988
C309	0.692	0.769	0.729	0.974	0.995

Indeed, how to balance the weights of ECs is one critical step in customer driven product design by employing QFD. Conventionally, customer survey data are collected and analyzed, which is often laborious. On the other hand, online reviews provide valuable information about CRs. However, these data are different with

For Peer Review Only

customer survey data. The effort in this research is on the analysis of online reviews to discover important ECs. An investigation is carried on to show that some methods of regression, classification and learning to rank are not capable to handle this problem. Accordingly, an ordinal classification approach is proposed to determine the weights of ECs by taking the overall customer satisfaction and the opinions of ECs in online reviews into considerations. Generally, it highlights which ECs are more important and required to be improved when designers are conceiving to launch new products. Moreover, the weights of ECs based on online reviews can be utilized in QFD and many other methods to examine customer concerns. Hence, without doubt, the analysis of online reviews will profit designers to understand CRs efficiently and effectively, which is extremely important in customer-driven product design.

6. CONCLUSIONS

The focus of this research is how to prioritize ECs based on online reviews. It is different from many research studies in the design area, which take conventional survey data as customer needs only. The needs in online reviews are assembled into product design directly to suggest the weights of ECs, which are one critical step in QFD.

In this research, limitations about simple classification methods and algorithms of learning to rank are analyzed at first. Accordingly, an ordinal classification approach is devised. It is a pairwise approach, in which the weights of ECs are learned from review pairs. Moreover, an integer linear programming problem

For Peer Review Only

is formulated in order to transform the pairwise results to the original levels of customer satisfaction, which is not explained in many pairwise approaches. In order to show how the proposed approaches benefit designers, an exploratory case study is carried on. In this case study, valuable review data labeled are gained, which contributes to initialize a sound analysis about customer reviews. Finally, based on these annotated review data, several categories of experimental studies are conducted and interesting phenomena are found.

For the future, there are many promising applications about how to exploit the value of online reviews from the perspective of product designers, such as, how to make comparison between different products based on online reviews, how to generate the quality of house directly based on online reviews, etc.

REFERENCES

Chris Burges, Tal Shaked, Erin Renshaw, Ari Lazier, Matt Deeds, Nicole Hamilton, and Greg Hullender. Learning to rank using gradient descent. In *ICML'05*. ACM, 2005.

L.-K. Chan and Wu, M.-L. Wu. A systematic approach to quality function deployment with a full illustrative example. *Omega*, 33, 119 - 139, 2005

Ankur Chaudha, Rajeev Jain, A. Singh, and P. Mishra. Integration of Kano's model into Quality Function Deployment (QFD). *The International Journal of Advanced Manufacturing Technology*, 53:689--698, 2011.

K. Y. Chan, C.K. Kwong and T.C.Wong. Modelling customer satisfaction for product development using genetic programming. *Journal of Engineering Design*, 22(1):55–68, 2011

Y. Chen, Richard Y. K. Fung, and Jiafu Tang. Rating technical attributes in fuzzy QFD by integrating fuzzy weighted average method and fuzzy expected value operator. *European Journal of Operational Research*, 174(3):1553--1566, 2006.

Reinhold Decker and Michael Trusov. Estimating aggregate consumer preferences from online product reviews. International Journal of Research in Marketing, 27(4): 293 -- 307, 2010

Yoav Freund, Raj Lyer, Robert E. Schapire, and Yoram Singer. An efficient boosting

For Peer Review Only

algorithm for combining preferences. *The Journal of Machine Learning Research*, 4:933--969, 2003.

Xiuli Geng, Xuening Chu, Deyi Xue, and Zaifang Zhang. An integrated approach for rating engineering characteristics' final importance in product-service system development. *Computers and Industrial Engineering*, 59(4):585 -- 594, 2010.

Saurabh Goorha and Lyle Ungar. Discovery of significant emerging trends. In *KDD'10*, pages 57-64, 2010

Ahmed Hassan and Dragomir Radev. Identifying text polarity using random walks. In *ACL'10*, pages 395--403, 2010

Ralf Herbrich, Thore Graepel, and Klaus Obermayer. *Large margin rank boundaries for ordinal regression*, page 115–132. MIT Press, Cambridge, MA, 2000.

Minqing Hu and Bing Liu. Mining and summarizing customer reviews. In *KDD'04*, pages 168--177, 2004

Thorsten Joachims. Optimizing search engines using clickthrough data. In *KDD'02*, pages 133--142, 2002

Thorsten Joachims, Hang Li, Tie-Yan Liu, and Chengxiang Zhai. Learning to rank for information retrieval (LR4IR 2007). *SIGIR Forum*, 41:58--62, December 2007.

C. Kahraman, T. Ertay & G. Buyukozkan. A fuzzy optimization model for QFD planning process using analytic network approach. *European Journal of Operational Research*, 171, 390 - 411, 2006

Soo-Min Kim and Eduard Hovy. Identifying opinion holders for question answering in opinion texts. In *AAAI'05 Workshop: Question Answering in Restricted Domains*, pages 20--26, 2005.

C. K. Kwong, Y. Chen, K.Y. Chan and X. Luo. A generalised fuzzy least-squares regression approach to modelling relationships in QFD. *Journal of Engineering Design*, 21(5): 601-613, 2010

C. K. Kwong, Y. Ye, Y. Chen, and K. L. Choy. A novel fuzzy group decision-making approach to prioritizing engineering characteristics in QFD under uncertainties. *International Journal of Production Research*, 49(19):5801--5820, 2011.

C. K. Kwong, T. C. Wong and K. Y. Chan. A methodology of generating customer satisfaction models for new product development using a neuro-fuzzy approach. *Expert Systems With Applications*, 36(8):11262-11270, 2009

Binyang Li, Lanjun Zhou, Shi Feng, and Kam-Fai Wong. A unified graph model for sentence-based opinion retrieval. In *ACL'10*, pages 1367--1375, 2010.

Zhaoling Li, Dongling Zhang, and Qisheng Gao. A grey method of prioritizing engineering characteristics in QFD. In *Control and Decision Conference*, pages 3443 --3447, June 2009.

Chenghua Lin and Yulan He. Joint sentiment/topic model for sentiment analysis. In *CIKM'09*, pages 375--384, 2009

Bing Liu, Minqing Hu, and Junsheng Cheng. Opinion observer: Analyzing and comparing opinions on the web. In *WWW'05*. ACM, 2005.

Tie-Yan Liu. Learning to Rank for Information Retrieval. In *SIGI*R'08. ACM, 2008.

Ying Liu, Wen Feng Lu, and Han Tong Loh. Knowledge discovery and management for product design through text mining - a case study of online information integration

For Peer Review Only

for designers. In *ICED'07*, pages 329/1--12, 2007

Ying Liu, Jian Jin, Ping Ji, Jenny A. Harding, and Richard Y. K. Fung. Identifying helpful online reviews: A product designer's perspective. *Computer-Aided Design*, 45(2):180-194, 2013

Ying Liu, Wen Feng Lu, and Han Tong Loh. Knowledge discovery and management for product design through text mining - a case study of online information integration for designers. In *Proceedings of the 16th international Conference on Engineering Design*, pages 329/1--12, 2007.

Tao Qin, Xu-Dong Zhang, De-Sheng Wang, Tie-Yan Liu, Wei Lai, and Hang Li. Ranking with multiple hyperplanes. In *SIGIR'07*, pages 279--286, 2007.

Guang Qiu, Bing Liu, Jiajun Bu, and Chun Chen. Expanding domain sentiment lexicon through double propagation. In *IJCAI'09*, pages 1199--1204, 2009.

Leonardo Rigutini, Tiziano Papini, Marco Maggini, and Franco Scarselli. Sortnet: Learning to rank by a neural preference function. *IEEE Transactions on Neural Networks*, 22:1368--1380, 2011.

Z. Sener and E. E. Karsak. A combined fuzzy linear regression and fuzzy multiple objective programming approach for setting target levels in quality function deployment. *Expert Systems with Applications*, 38, 3015 - 3022, 2011

Conrad Tucker and Harrison M. Kim. Predicting emerging product design trend by mining publicity available customer review data. In *ICED'11*, pages 43-52, 2011

J. Wang. Fuzzy outranking approach to prioritize design requirements in quality function deployment. *International Journal of Production Research*, 37, 899-916, 1999

L. Wang, B. D. Youn, S. Azarm and P. K. Kannan. Customer-driven product design selection using web based user-generated content. In *IDETC'11*, pages 405-419, 2011

Ying-Ming Wang and Kwai-Sang Chin. Technical importance ratings in fuzzy QFD by integrating fuzzy normalization and fuzzy weighted average. *Computers and Mathematics with Applications*, 62(11):4207 -- 4221, 2011.

Taras Zagibalov and John Carroll. Automatic seed word selection for unsupervised sentiment classification of Chinese text. In *COLING'08*, pages 1073--1080. Association for Computational Linguistics, 2008.

Jiaming Zhan, Han Tong Loh, and Ying Liu. Gather customer concerns from online product reviews - a text summarization approach. *Expert Systems With Applications*, 36(2 Part 1):2107-- 2115, 2009.

Wei Zhang, Lifeng Jia, Clement Yu, and Weiyi Meng. Improve the effectiveness of the opinion retrieval and opinion polarity classification. In *CIKM'08*, pages 1415--1416, 2008.

For Peer Review Only

Figure Captions

1,651 of 1,738 people found the following review helpful

★★★☆☆ **Good, but ...**, March 25, 2012

By **Bookenz**

This review is from: Apple iPad MD328LL/A (16GB, Wi-Fi, White) 3rd Generation (Personal Computers)

I've been a big iPad fan and was waiting anxiously for the third generation to come out. I bought the original iPad, and when the 2 came out, I happily sold the 1 and upgraded. I was thrilled with the thinner, lighter and improved iPad 2. So naturally, when word came out that the third one was about to be released in March 2012, I was right on board to buy one. Sold my 2 and preordered the 3 from Apple the day it was released.

I was hoping the third gen wouldn't be noticibly thicker and heavier than the 2, but unfortunately it was. I could definitely tell the difference when reading ebooks, which I do a lot. I couldn't really tell any difference between the speed and clarity of the 3, but to be fair, I didn't compare the two models side by side. I've no doubt the 3 is superior in this regard. I don't use the cameras, so don't care about this since I have a very nice digital camera for that.

The one thing about the third generation iPad that concerned me was the heat issue. Shortly after receiving it, I was reading an ebook and noticed the left side was warm. Not hot, but definitely warm enough for me to notice it. This reminded me of laptops I've had that have overheated and shut down, and here I was only reading a book. My iPad 2 never had this problem and I used it a lot. I also had some difficulty backing up the 3 to the cloud, again, something that wasn't ever a problem with the 2.

After reading some reviews of others experiencing the heating problem with the latest iPad, and really missing the thinner and lighter iPad 2 , I decided to return the iPad 3 to my local Apple store and buy a new iPad 2. They had no problem taking it back and I was glad to see the 2 had come down in price. When the sales associate asked me why I was returning the 3, I told him about the heat problem. He didn't seem surprised and said it was because the 3 has a larger battery.

The third gen has a faster CPU and retina dsplay, but I never thought the 2 had any problems with speed, and the clarity of the display has always seemed fine to me. But I'm certainly no expert. I'm very happy I made the exchange and enjoying using my iPad 2 again. For my uses--books, internet, email and games, the cooler, thinner and lighter iPad 2 is a better choice for me.

By the way, I just want to add that I think it's unfair to compare the Kindle Fire and Nook to the iPad. These devices are just ebook readers and they're fine for that--I've tried them both. Rather like saying, okay, you can buy this Mercedes or you can get the Volkswagen. Once you've used an iPad, you're totally spoiled and trying to browse the Internet with a Kindle or Nook is an exercise in frustration after using an iPad.

Update January 7, 2014 - since writing this review for the iPad 3, which I ended up returning, I've moved on to the iPad Air, which is noticeably lighter and faster than the iPad 2.

Help other customers find the most helpful reviews | Report abuse | Permalink

Was this review helpful to you? Yes No

Figure 1. One typical online review

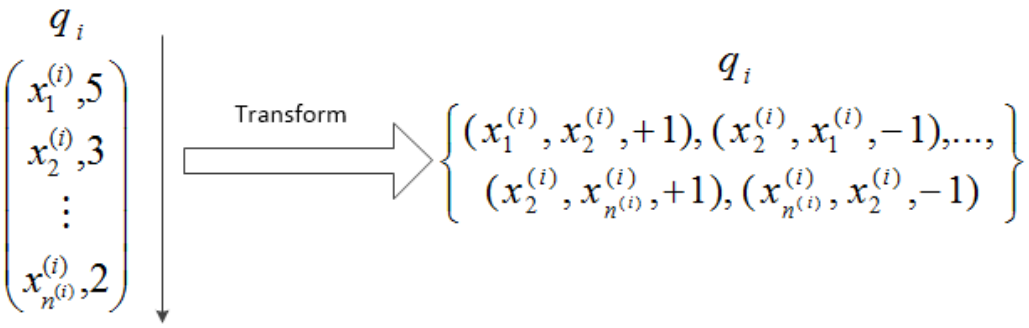


Figure 2. An conceptual illustration for pairwise approaches of learning to rank (Liu, 2008)

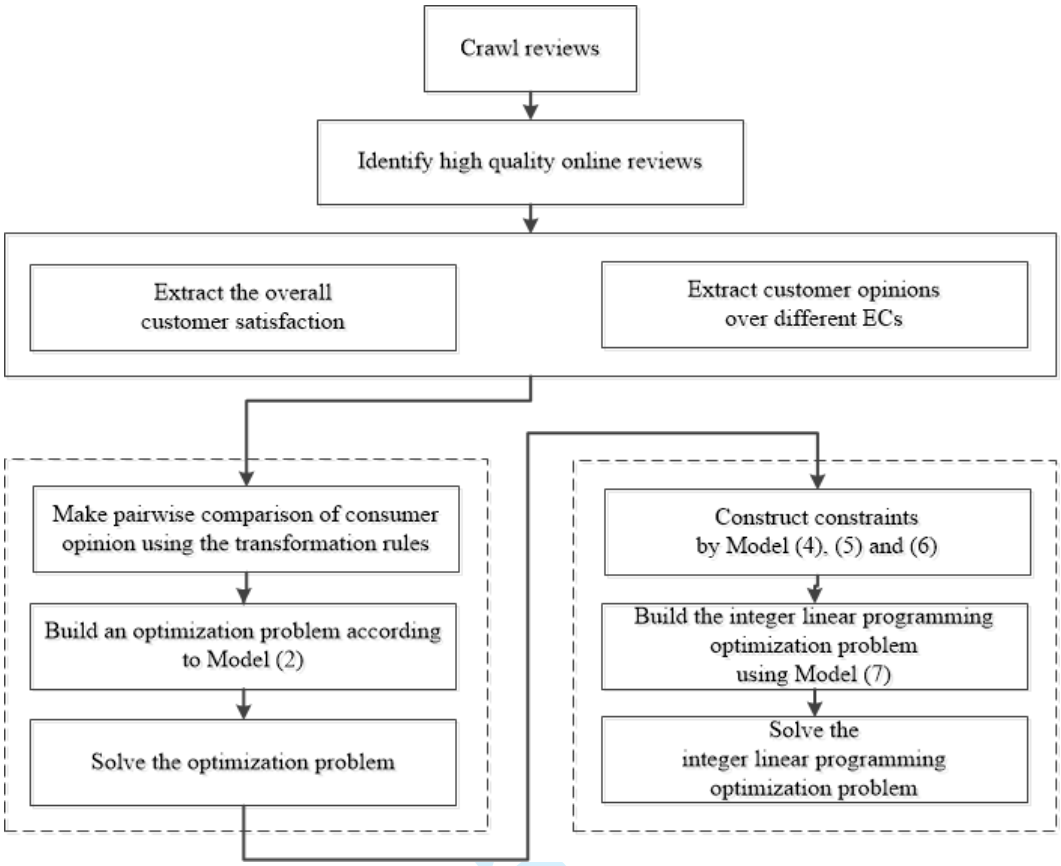


Figure 3. the procedure of prioritizing ECs from online reviews

For Review Only

Sentence		Auto Document Feeder	Card Slot	Consumable Replacement	Durability	Ease of setup	Ease of use	Fax Setting	Hopper Unit	Mac Compatible	Ink Longevity	LCD Panel	Noise	Power Supply	Print Quality	Printer Housing	Wifi Integration
I only have the printer for a few days, but so far I am very pleased.	-																
I was a bit nervous before buying because I had read a lot of complai	-																
But the paper handling on my 810 has been flawless, if a tiny bit noi	-												-2				
But the paper handling on my 810 has been flawless, if a tiny bit noi	-								-1								
The actual printing is quiet, and of great quality.	-												1				
The actual printing is quiet, and of great quality.	-														2		
The paper tray feels a bit flimsy, but is easy to remove or insert, a	-								-2								
It can expand to hold legal size paper, and has a separate area for s	-																
The package comes with 3 sheets of Epson high quality 4x6 glossy phot	-														2		
I have been using an HP Wireless printer up until now(model 5850)and	-														2		
Also, they promise that their inks are acid-free and will last 3 or 4	-										0						
Another pre-purchase worry was that there were complaints about the A	-																
Too early to tell, but I 've printed those 3 photos and a fourth on p	-										1						
So, so far so good on ink usage.	-										0						
I was surprised and pleased too that the 810 printer comes with an TW	-			1													
Nice!	-																
Additionally, the inks provided are the same capacity as the refills,	-			2													
A nice feature is that when you plug in a camera memory card in the f	-		2														
Very neat.	-																
It warns on the box that this feature may not work with Mac computers	-									1							
Have n 't tried printing from the iPhone as they claim you can do, tho	-																
I love the big touch screen operation too, it is pretty slick.	-											2					
Made a few copies of some simple pencil drawings and they were very g	-														-1		
Updated Review 03\/2010Here 's a 6 month update.	-																
The printer is getting a lot of work ; we print a lot of classroom ma	-				2												

Figure 4. Review labeling

Review Only

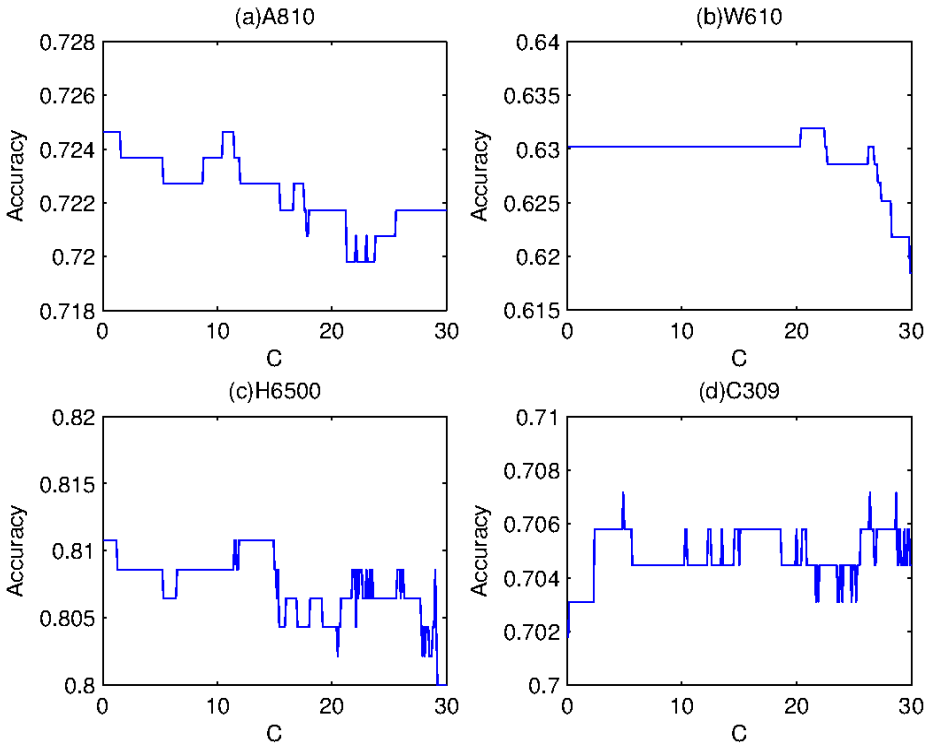


Figure 5. Accuracy vs. regularization term *C*

Review Only

1,651 of 1,738 people found the following review helpful

★★★☆☆ **Good, but ...**, March 25, 2012

By **Bookenz**

This review is from: **Apple iPad MD328LL/A (16GB, Wi-Fi, White) 3rd Generation (Personal Computers)**

I've been a big iPad fan and was waiting anxiously for the third generation to come out. I bought the original iPad, and when the 2 came out, I happily sold the 1 and upgraded. I was thrilled with the thinner, lighter and improved iPad 2. So naturally, when word came out that the third one was about to be released in March 2012, I was right on board to buy one. Sold my 2 and preordered the 3 from Apple the day it was released.

I was hoping the third gen wouldn't be noticibly thicker and heavier than the 2, but unfortunately it was. I could definitely tell the difference when reading ebooks, which I do a lot. I couldn't really tell any difference between the speed and clarity of the 3, but to be fair, I didn't compare the two models side by side. I've no doubt the 3 is superior in this regard. I don't use the cameras, so don't care about this since I have a very nice digital camera for that.

The one thing about the third generation iPad that concerned me was the heat issue. Shortly after receiving it, I was reading an ebook and noticed the left side was warm. Not hot, but definitely warm enough for me to notice it. This reminded me of laptops I've had that have overheated and shut down, and here I was only reading a book. My iPad 2 never had this problem and I used it a lot. I also had some difficulty backing up the 3 to the cloud, again, something that wasn't ever a problem with the 2.

After reading some reviews of others experiencing the heating problem with the latest iPad, and really missing the thinner and lighter iPad 2 , I decided to return the iPad 3 to my local Apple store and buy a new iPad 2. They had no problem taking it back and I was glad to see the 2 had come down in price. When the sales associate asked me why I was returning the 3, I told him about the heat problem. He didn't seem surprised and said it was because the 3 has a larger battery.

The third gen has a faster CPU and retina dsplay, but I never thought the 2 had any problems with speed, and the clarity of the display has always seemed fine to me. But I'm certainly no expert. I'm very happy I made the exchange and enjoying using my iPad 2 again. For my uses--books, internet, email and games, the cooler, thinner and lighter iPad 2 is a better choice for me.

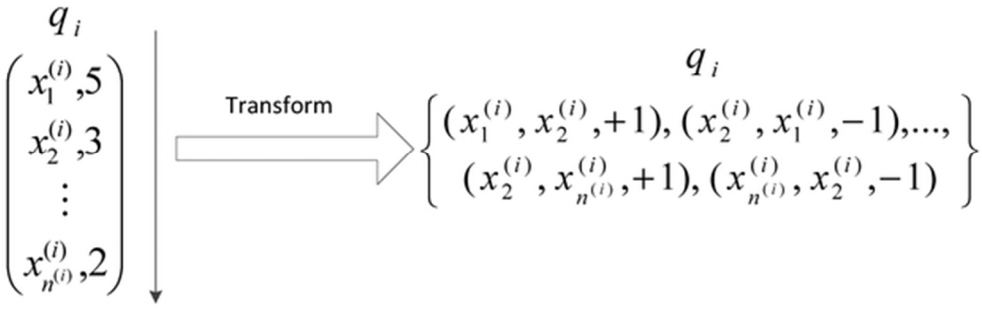
By the way, I just want to add that I think it's unfair to compare the Kindle Fire and Nook to the iPad. These devices are just ebook readers and they're fine for that--I've tried them both. Rather like saying, okay, you can buy this Mercedes or you can get the Volkswagen. Once you've used an iPad, you're totally spoiled and trying to browse the Internet with a Kindle or Nook is an exercise in frustration after using an iPad.

Update January 7, 2014 - since writing this review for the iPad 3, which I ended up returning, I've moved on to the iPad Air, which is noticeably lighter and faster than the iPad 2.

Help other customers find the most helpful reviews | Report abuse | Permalink

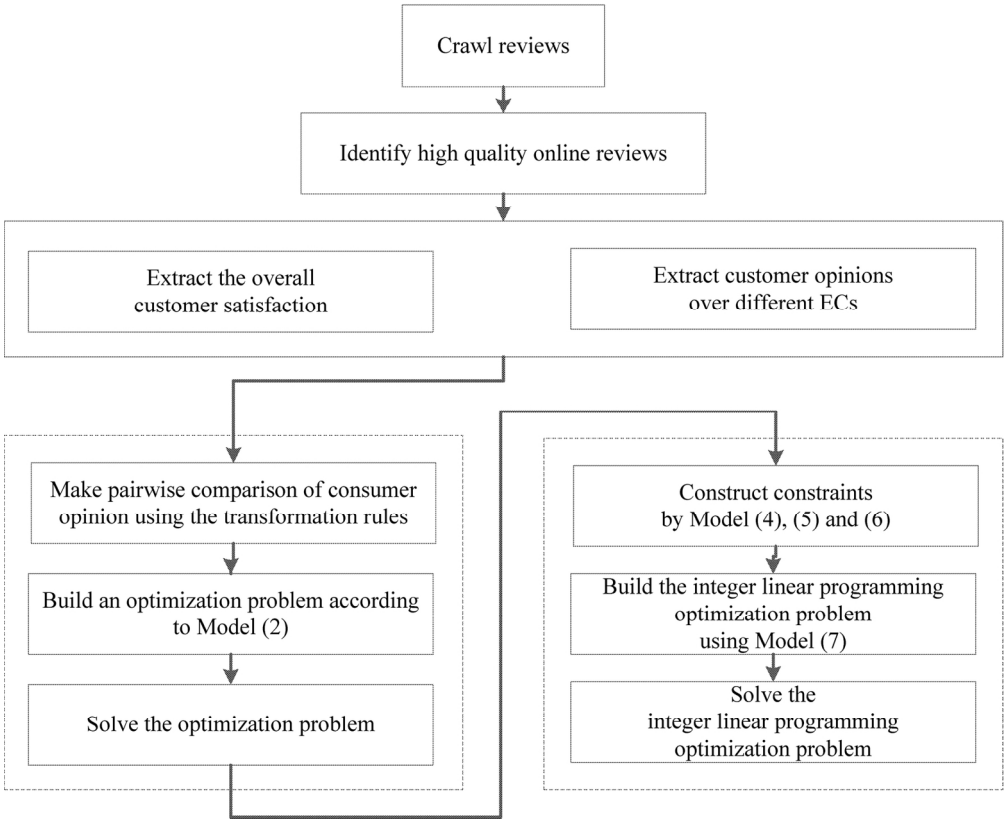
Was this review helpful to you? Yes No

152x188mm (300 x 300 DPI)



53x17mm (300 x 300 DPI)

Peer Review Only



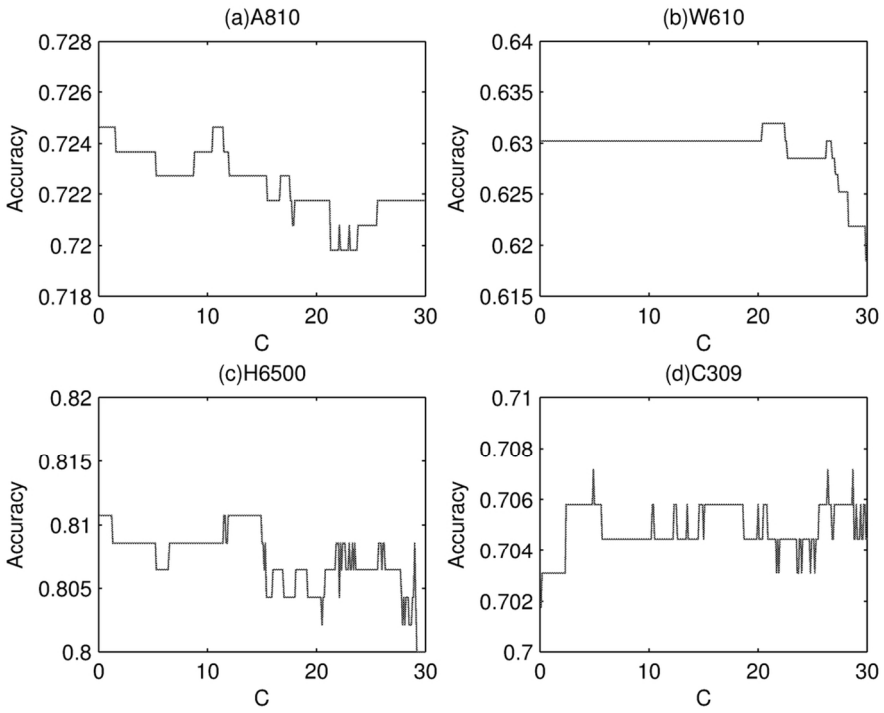
150x122mm (300 x 300 DPI)

For Peer Review Only

Sentence		Auto Document Feeder	Card Slot	Consumable Replacement	Durability	Ease of setup	Ease of use	Fax Setting	Hopper Unit	Mac Compatible	Ink Longevity	LCD Panel	Noise	Power Supply	Print Quality	Printer Housing	Wifi Integration
I only have the printer for a few days, but so far I am very pleased.	-																
I was a bit nervous before buying because I had read a lot of complai	-																
But the paper handling on my 810 has been flawless, if a tiny bit noi	-												-2				
But the paper handling on my 810 has been flawless, if a tiny bit noi	-								-1								
The actual printing is quiet, and of great quality.	-												1				
The actual printing is quiet, and of great quality.	-														2		
The paper tray feels a bit flimsy, but is easy to remove or insert, a	-								-2								
It can expand to hold legal size paper, and has a separate area for s	-																
The package comes with 3 sheets of Epson high quality 4x6 glossy phot	-														2		
I have been using an HP Wireless printer up until now(model 5850)and	-														2		
Also, they promise that their inks are acid-free and will last 3 or 4	-										0						
Another pre-purchase worry was that there were complaints about the A	-																
Too early to tell, but I 've printed those 3 photos and a fourth on p	-										1						
So, so far so good on ink usage.	-										0						
I was surprised and pleased too that the 810 printer comes with an TW	-			1													
Nice!	-																
Additionally, the inks provided are the same capacity as the refills,	-			2													
A nice feature is that when you plug in a camera memory card in the f	-		2														
Very neat.	-																
It warns on the box that this feature may not work with Mac computers	-									1							
Have n 't tried printing from the iPhone as they claim you can do, tho	-																
I love the big touch screen operation too, it is pretty slick.	-											2					
Made a few copies of some simple pencil drawings and they were very g	-														-1		
Updated Review 03\/2010Here 's a 6 month update.	-																
The printer is getting a lot of work ; we print a lot of classroom ma	-				2												

167x139mm (300 x 300 DPI)

ew Only



111x83mm (300 x 300 DPI)

iew Only