Decision Support Systems 48 (2010) 470–479

Contents lists available at ScienceDirect

Decision Support Systems

journal homepage: www.elsevier.com/locate/dss

Maximizing customer satisfaction through an online recommendation system: A novel associative classification model

Yuanchun Jiang [a,b,*], Jennifer Shang [b], Yezheng Liu [a,c]

[a] School of Management, Hefei University of Technology, Hefei, Anhui 230009, China
[b] The Joseph M. Katz Graduate School of Business, University of Pittsburgh, Pittsburgh, PA 15260, USA
[c] Key Laboratory of Process Optimization and Intelligent Decision Making, Ministry of Education, Hefei, Anhui 230009, China

ARTICLE INFO

Available online 17 June 2009

Keywords:
Online recommendation
Customer satisfaction
Associative classification
Rating classification

ABSTRACT

Offering online personalized recommendation services helps improve customer satisfaction. Conventionally, a recommendation system is considered as a success if clients purchase the recommended products. However, the act of purchasing itself does not guarantee satisfaction and a truly successful recommendation system should be one that maximizes the customer's after-use gratification. By employing an innovative associative classification method, we are able to predict a customer's ultimate pleasure. Based on customer's characteristics, a product will be recommended to the potential buyer if our model predicts his/her satisfaction level will be high. The feasibility of the proposed recommendation system is validated through laptop Inspiron 1525.

1. Introduction

Personalization of product information has become one of the most important factors that impact a customer's product selection and satisfaction in today's competitive and challenging market. Personalized service requires firms to understand customers and offer goods or services that meet their needs. Successful firms are those that provide the right products to the right customers at the right time and for the right price.

As a type of information technology aimed to support personalized service, recommendation systems are widely used by e-commerce practitioners and have become an important research topic in information sciences and decision support systems [25]. Recommendation systems are decision aids that analyze customer's prior online behavior and present information on products to match customer's preferences. Through analyzing the patron's purchase history or communicating with them, recommendation systems employ quantitative and qualitative methods to discover the products that best suit the customer. Most of the current recommendation systems recommend products that have a high probability of being purchased [3]. They employ content-based filtering (CBF) [41], collaborative filtering (CF) [18], and other data mining techniques, for example, decision tree [12], association rule [38], and semantic approach [25]. Other literature focuses on the influence of recommendation systems on customer's purchase behavior [3,32]. They argue that the recommendation decision should be based not on purchase probability, but rather on the sensitivity of purchase probability due to the recommendation action. Common wisdom regards a recommendation system as successful if customers end up purchasing the suggested product(s). However, buying a product does not necessarily imply the client is pleased with the product. Let's consider a scenario below.

James is in need of a laptop computer. He visits online stores to look for information and compare prices and performance of various laptops. Between the two laptop series, Inspiron 1525 and Aspire 5735, James is uncertain which one would best fit his needs. He decides to turn to the recommendation system for help. After gaining knowledge of James's needs and personal profile, the system recommends the Inspiron 1525. Once James follows the advice and makes his purchase, the recommendation system deems that it did a great job because James bought the laptop it recommended. However, after 1 week's use of the laptop, James writes a review as follows: "…a good product, but not the one I really want." It turns out James is not content with the recommendation. This exemplifies the case that a customer may have purchased the recommended product(s), but the recommendation system was not successful in pleasing the customer—its ultimate goal. It is therefore clear that a customer's acceptance of a recommendation is not equivalent to its success. A recommendation system must endure the test of time. Only when customers claim that the products are what they like after their practical usage can one claim that the system has made effective recommendations. This requires not only matching customers' needs, but also satisfying customers' wants. In other words, the recommendation system should only recommend a product if its satisfaction rating is predicted to be high.

How can a customer's satisfaction of a specific product be measured and attained? The rapid development of e-commerce affords us an opportunity to predict customers' reactions after they use a product. Many online stores, such as Amazon.com and Dell.com encourage

* Corresponding author. School of Management, Hefei University of Technology, Hefei, Anhui 230009, China.
E-mail address: yuanchunjiang@gmail.com (Y. Jiang).

0167-9236/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.dss.2009.06.006

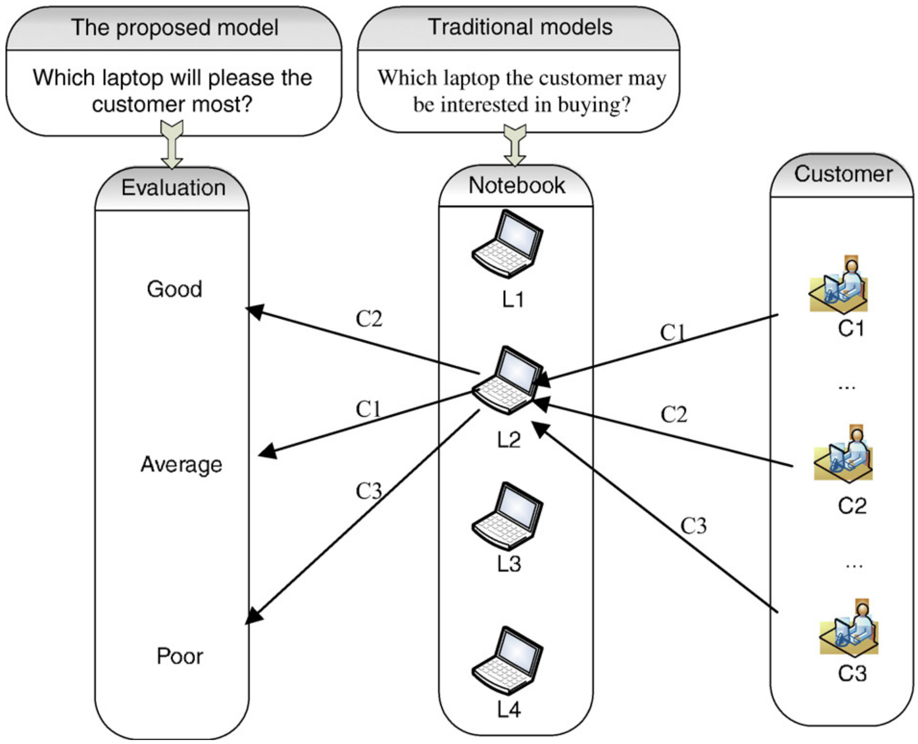

Fig. 1. Differences between the existing recommendation systems and the proposed model.

customers to write online reviews on their websites; information from these reviews is then often used to support a firm's product strategy and customer relationship management [11,13]. In the online reviews, customers can discuss their needs, preferences, personal profile, and voice their opinions about a product as poor, average, or good. From such need-rating data, it is easy to obtain personalized information and customers' after-use satisfaction level of the product. Using personal information and responses, the online store can more accurately predict customers' true sentiments toward a specific product, and recommend a more suitable product for the potential customer to enjoy.

This research proposes a rating classification model to estimate a potential customer's satisfaction level. It builds a rating classifier for a product by discovering rules from the need-rating database collected for the product. The rules imply the co-relationship between customers' needs, preferences, demographic profile, and their ratings for the product. For a new customer with specific characteristics, the classifier will predict his/her response toward the recommended product and categorize it into certain class labels, such as poor, average and good. The predicted ratings estimate the customer's satisfaction level for the product. Differences between the existing recommendation systems and the proposed one are illustrated in Fig. 1.

This research proposes a novel associative classification model, which first mines multi-class classification information from need-rating data, then constructs a rating classifier, and finally predicts customers' ratings for products. We organize the rest of the paper as follows. In Section 2 we review the literature of recommendation systems and associative classification models. Section 3 proposes the innovative methodology to address the rating classification problem. A case study used to illustrate the effectiveness of the proposed model is given in Section 4. Section 3 comprises the Summary, conclusions, and future research.

2. Literature review

The literature review focuses on two perspectives: the recommendation system and associative classification. A summary of relative research methods are given in Table 1 and explained in detail below.

2.1. Recommendation systems

Since the development of the first recommendation system by Goldberg and colleagues [17], various recommendation systems and related technologies such as CBF and CF [18,41] have been reported. Among them, the user-based collaborative filtering (CF) [23] is successfully adopted by Amazon.com and Dell.com. It finds a similar user group for the target buyer and recommends products that have been rated by users in the reference group but not yet viewed by the target buyer. However, the user-based CF has some limitations. One is its difficulty in measuring the similarities between users, and the other is the scalability issue. As the number of customers and products increases, the computation time of algorithms grows exponentially [21]. The item-based CF [16] was proposed to overcome the scalability problem as it calculates item similarities in an offline basis. It assumes that a user will be more likely to purchase items that are similar or related to the items that he/she has already purchased.

Table 1
Summary of research methods on recommendation system and associative classification.

(a) Motivation and objectives of various recommendation systems

Literature	Motivation	Objective
[2,10,16,17,20,21,23,41]	Which products meet the customer's preferences best?	Recommend products with high probability.
[3,32]	What is the influence of recommendation systems on customer's purchase behavior?	Recommend products which are receptive to the recommendation.
This paper	Which products can achieve a high after-use satisfaction level?	Recommend products with high after-use satisfaction level.

(b) Comparing associative classification models

Literature	Mine multi-class rules	Classify using multiple rules	Provide classification reasons
[26,31,36,37]	√		
[24]		√	
[27] and this paper	√	√	√

The content-based filtering (CBF) method applies content analysis to target items. Target items are described by their attributes, such as color, shape, and material. The user's profile is constructed by analyzing his/her responses to questionnaires, his/her rating of products, and navigation history. The recommendation system proposes items that have high correlations with a user's profile. However, a pure CBF system also has its shortcomings. One is that users can only receive recommendations similar to their earlier experiences. The other is that some items, such as music, photographs, and multimedia, are hard to analyze [10]. Based on CF and CBF, new data mining techniques employing decision tree, association rule, regression model, and Markov chain have been introduced to recommend movies and books [2], support one-to-one online marketing [21], and attract customers for the tourism industry [20].

Unlike those recommending products based on likelihood of purchase, Bodapati [3] argued that the recommendation decision should also examine a customer's sensitivity to such a recommendation. He built a model to measure the role of recommendation systems in modifying customers' purchase behavior relative to what the customers would have done without such recommendation interventions. Although the extant recommendation systems may recommend acceptable products to customers, they share a common view that the act of purchase itself equates to the customers' satisfaction, which could be far from the truth, as evidenced by James' example earlier.

2.2. Associative classification and the combination strategy for multi-class classification

Classification is an important management task. Many methods such as the agent-based approach [34], decision tree [30], and data envelopment analysis (developed by professor Cooper [14]) have been proposed to solve the decision analysis problems in various fields. Associative classification is a relatively new classification method whose aim is to apply the Apriori algorithm [1] to mine association rules and construct associative classifiers [26]. Rule mining will find the associations between attributes (rule preconditions) and ratings (results). In associative classification, the support degree is defined as the ratio of the number of objects satisfying a specific rule precondition and having a specific rating result over the total number of objects in the database. The confidence degree is similar to the support degree except that the number of all objects satisfying the specific rule precondition is used as the denominator. The discovered rules are pruned to attain a minimal rule set necessary to cover training data and achieve sufficient accuracy [37]. Although associative classification methods may derive more accurate classification results than other methods, they have a few drawbacks [27]. First is related to multi-class classification: The associative classification methods available today do not have enough multi-class information to build multi-class classifiers because all conflicting rules are removed [31]. For example, $P_1 \rightarrow c_1$ and $P_1 \rightarrow c_2$ are two conflicting rules having the same precondition P_1 but different classifications, c_1 and c_2, with confidence degrees of 51% and 49%, respectively. Traditional associative classification methods will delete rule $P_1 \rightarrow c_2$ because its confidence level is lower than that of $P_1 \rightarrow c_1$. When conflicts occur, only the rule with the highest confidence level is retained; the competing rules having lower probabilities are all removed. Another flaw is that they cannot easily identify an optimal rule when classifying a new case [24]. An optimal rule is the one that maximizes the measure, such as the support degree, confidence degree, or interesting degree as defined by the user. The difficulty with the traditional methods is that different measures may result in different optimal rules.

To overcome the above weakness, Liu and his colleagues [27] have proposed a combination strategy for multi-class classification (CSMC). CSMC retains most of the conflicting rules and employs multiple association rules to construct classifiers for new cases. After acquiring conflicting rules and calculating their weights, the evidential reasoning approach proposed by Yang and colleagues [40] is employed to combine the classification results of distinct rules. CSMC is useful since it applies multiple rules—especially conflicting ones—to multi-class classification. However, it has deficiencies as well. That is, even though CSMC retains multi-class information in the process of rule acquisition, much of it is lost in rule pruning. Consequently, the evidence bodies used in the rating classification are often inaccurate and classification results suffer. Moreover, CSMC employs a rough set method to derive evidence weights. Although using evidence weights may significantly improve classification accuracy, its computation is very time-consuming and negatively affects the efficiency of classification.

In this research, we propose a new algorithm to address the rating classification problem. Compared with CSMC, the proposed algorithm can preserve most of the useful multi-class information after pruning and it can derive attribute weights much more efficiently. Details of our algorithm are discussed next.

3. The proposed methodology

3.1. The solution framework

To construct an efficient and powerful rating classifier for a specific product, we follow three phases as outlined in Fig. 2. The first phase is to mine need-rating rules, where retaining multi-class information is the main task. Since conflicts and ambiguities are often present in the need-rating database, the rules discovered must be able to deal with contradictory facts and uncertainties, and to arrive at multi-class information. The second phase calculates the weight for each rule. A weight measures the importance of a need-rating rule in the rating classification problem. In this phase, computational efficiency is crucial, due to the presence of various rules useful for classifier construction. The last phase is to develop the rating classifier, whose goal is to recommend products that yield high customer satisfaction. In this phase, all factors having influences on the potential customers' satisfaction levels are considered. Since customers of the same need may voice very different opinions for the same product, it is important to make multi-class ratings available. Predicting after-use ratings along with corresponding likelihoods provides potential customers a valuable purchase guideline, which significantly enhances the odds of customer satisfaction.

Although traditional associative classification methods could generate reasonable classification results, they suffer from three weak points when classifying customers' ratings. First, they do not accommodate conflicting ratings. Customers of the same needs and preferences may have very different opinions (ratings) for the same product. As discussed in Section 3, traditional associative classification methods resolve the situation by retaining only the rule that has the highest confidence level. As a result, not enough multi-class information is preserved to deal with the conflict nature. The second weakness is the prediction tactic. To predict accurately, a classifier needs to consider a customer's needs, preferences, and demographics. However, traditional prediction relies only on one optimal rule, which is not enough to attain accurate classification. Thirdly, traditional methods classify each case without explanations, that is, they simply list the applicable rule. Such classification is somewhat arbitrary, ambiguous, and irrational. Since customers of the same need may give different ratings to the same product, a classification method must be capable of predicting the probabilities of attaining different customer ratings. The proposed classification algorithm aims to overcome the above deficiency. Details are described next.

3.2. Mine need-rating rules

The need-rating data is first organized into a data table $I = (O, A \cup C)$, where O is the set of objects (customers), $|O|$ is the number of customers in I. A is the set of attributes, $A = \{A_1, \ldots, A_h, \ldots, A_{|A|}\}$, $|A|$ is the number of attributes in A, each A_h, $h = 1,2,\ldots,|A|$, is a factor/criterion or a customer characteristic. C corresponds to a set of class labels, $C = \{c_1, \ldots, c_g, \ldots, c_{|C|}\}$, $|C|$ is the number of classes in I, each c_g, $g = 1,2,\ldots,|C|$, denotes a rating grade.

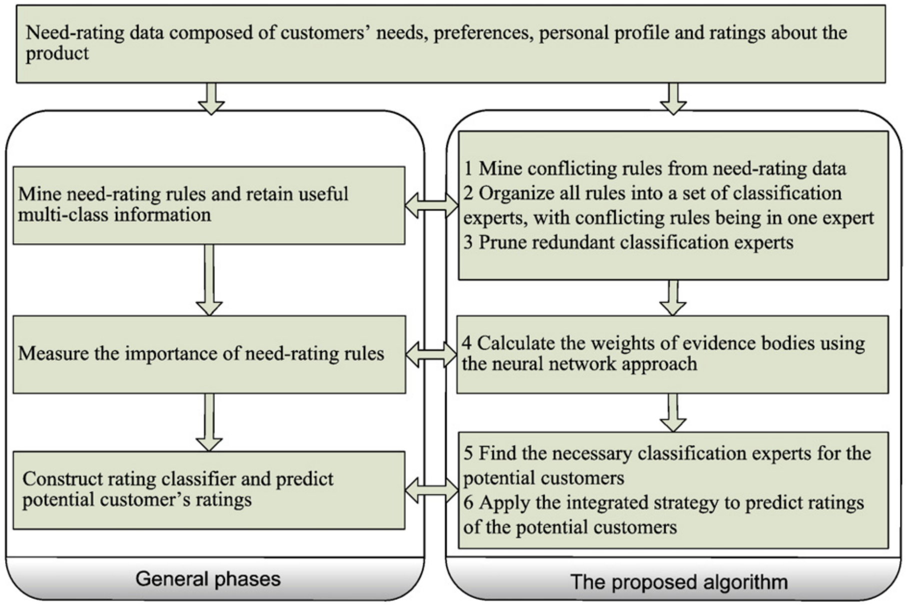

Fig. 2. The proposed rating classification framework.

In data table I, customers are described by customer characteristics, preferences, and ratings for the product. For example, a middle-aged customer who prefers a laptop with average central processing unit (CPU) speed, good battery life, and rates the laptop Inspiron 1525 as a good product can be described as follows:

$$('Age = Middle' \wedge 'CPU = Average' \wedge 'Battery = Good') \wedge 'Rating = Good'.$$

The rule mining algorithm first scans the data table once to mine need-rating rules involving one attribute in rule preconditions. It then recursively combines the need-rating rules generated to extract rules involving more attributes. The support and confidence degrees for rules are calculated simultaneously. Any rules with support and confidence degrees larger than the threshold are saved as need-rating rules. For example, $P_1 \rightarrow c_1$, $P_1 \rightarrow c_2$, and $P_1 \rightarrow c_3$ are three conflicting rules with the same precondition. Their confidence degrees are 49%, 43%, and 8%, respectively. If the minimal confidence threshold is 40%, then rules $P_1 \rightarrow c_1$ and $P_1 \rightarrow c_2$ would be retained, whereas $P_1 \rightarrow c_3$ is pruned.

After a comprehensive set of rules are mined, the redundant ones are deleted. The proposed prune process eliminates redundancy to ensure that succinct need-rating rules are derived, and all necessary multi-class information is preserved.

3.2.1. Derivation of classification experts

Let R be the set of need-rating rules discovered from data table I:

$$R = \{P_1 \rightarrow c_1, \cdots, P_d \rightarrow c_d, \cdots, P_{|R|} \rightarrow c_{|R|}\}$$

where $|R|$ is the number of rules in R. Precondition P_d consists of the attribute-values in data table I, and result c_d is the corresponding rating. For any subset R_i of the rule set R,

$$R_i = \{P_{i,1} \rightarrow c_{i,1}, \cdots, P_{i,m} \rightarrow c_{i,m}, \cdots, P_{i,|R_i|} \rightarrow c_{i,|R_i|}\}$$

where $|R_i|$ is the number of rules in R_i, $P_{i,m} \rightarrow c_{i,m}$ is the mth rule in R_i. R_i is a classification expert in the rating classification problem if the following constraints are satisfied:

(1) $P_{i,1} = \cdots = P_{i,m} = \cdots P_{i,|R_i|}$
(2) For any other rules $(P_d \rightarrow c_d) \notin R_i$, $P_d \neq P_{i,m}$
(3) $c_{i,1} \neq \cdots \neq c_{i,m} \neq \cdots \neq c_{i,|R_i|}$.

R_i consists of all the need-rating rules which have the same precondition but different ratings. That is, it includes all the multi-class information associated with $P_{i,m}$. For convenience, we rewrite R_i as: R_i: $P_i \rightarrow E_i$, where

$$P_i = P_{i,m}, E_i = (c_{i,1}, conf_{i,1}) \vee \cdots \vee (c_{i,m}, conf_{i,m}) \vee \cdots \vee (c_{i,|R_i|}, conf_{i,|R_i|}) \vee (\Theta, conf_{i,\Theta})$$

$\mathrm{conf}_{i,m}$ is the confidence degree of $P_{i,m} \rightarrow c_{i,m}$, which is larger than the minimal confidence threshold, and Θ is the frame of discernment defined in the evidence theory. The $\mathrm{conf}_{i,\Theta}$ is the belief degree assigned to catchall (default) terms since they cannot be classified to any specific ratings by rules in R_i:

$$conf_{i,\Theta} = 1 - \sum_{m=1}^{|R_i|} conf_{i,m}$$

In evidence theory, evidence bodies are given by experts and consist of hypothesis–probability pairs. For example, the evidence body may look like: {(Good, 0.6), (Average, 0.3), (poor, 0.1)}. In the expert's opinion, the probability of receiving a good evaluation is 60%, an average evaluation is 30%, and a poor evaluation is 10%. Recall that $\mathrm{conf}_{i,m}$ can be treated as a belief degree given by expert R_i to a hypothesized rating $c_{i,m}$, based on observed attribute values in P_i. Therefore, E_i is regarded as the evidence body provided by classification expert R_i.

The set of need-rating rules is transformed into a set of classification experts, $R = \{R_1, \ldots, R_i, \ldots, R_j, \ldots, R_T\}$, which satisfy the following constraints:

(1) $\bigcup_{i=1}^{T} R_i = R$

(2) $R_i \cap R_j = \varphi$, for any i and j, $i \neq j$.

The set of evidence bodies corresponding to R is denoted as $E = \{E_1, \ldots, E_i, \ldots, E_j, \ldots, E_T\}$. To this point, we have transformed all the need-rating rules into T independent classification experts. However, not all classification experts are necessary for the recommendation system; some are redundant. In the next subsection, we develop pruning methods to remove the redundant classification experts.

3.2.2. Classification expert pruning through an improved algorithm

The goal of pruning classification experts is to generate a minimal set of classification experts which can cover all customers in data table *I*. Before presenting the pruning algorithm, we define the order of a relationship on classification experts.

For two classification experts R_i and R_j, R_j is said to be inferior to R_i, that is, $R_i \succ R_j$, if

1. $P_i \subset P_j$
2. For $\forall\ (c_{j,n}, \mathrm{conf}_{j,n}) \in E_j$, $\exists\ (c_{i,m}, \mathrm{conf}_{i,m}) \in E_i$, $c_{i,m}$ is the same as $c_{j,n}$ and $\mathrm{conf}_{j,n} = \mathrm{conf}_{i,m}$.

The relationship $R_i \succ R_j$ implies that classification experts R_i and R_j provide the same classification outcome, but R_i gives the classification information based on a simpler precondition. The classification expert R_j is more restrictive but not more powerful relative to R_i. Hence, R_j is redundant and should be removed to improve the quality of the classifier. Fig. 3 shows the pruning algorithm used to improve the quality of classification experts.

The first step prunes inferior classification experts. A classification expert R_j should be pruned if $\exists R_i \in \mathrm{CES}$, $R_i \succ R_j$. That is, if the classification expert has more complex preconditions but does not provide added information, then delete it. A simpler precondition is favorable since it provides a more powerful classification capability. The next pruning step uses the data covering method [26]. Classification expert R_s is necessary for customer *c* if R_s is a matching classification expert for *c* and there is not another matching classification expert whose precondition includes that of R_s. Furthermore, we remove the classification experts which do not meet the support or confidence threshold.

3.3. Measure the importance of evidence bodies

After classification experts are generated, the next step is to determine the weights of the evidence bodies given by classification experts. In the traditional evidence theory, evidence weights are consistent with human experts' knowledge and experience. If a human expert is authoritative and familiar with the decision problem, the evidence body given by her/him will be reliable. In this research, the evidence bodies are derived from different attribute sets. Thus we use the weights of attributes to infer the importance of evidence bodies.

```
Input: The classification expert set (R) and data table (I)
Output: Final classification expert set (CES)

The pruning algorithm:
  CES←R
 1. For each classification expert Rj in CES
      if ∃Ri ∈ CES , Ri≻Rj
              CES = CES \ Rj
      end if
    End for

 2. For each classification expert Rs in CES
       counts←0
     For each customer c in data table I
      If Rs is a necessary for c
        counts←counts+1
      End if
     End for
    End for
    Delete the classification experts in CES which
 dissatisfy the following constraint:
                     counts ≥ threshold
```

Fig. 3. The classification expert pruning algorithm.

Many methods such as support vector machine (SVM) [33] and information theory [9] have been used to measure attribute weights. Neural network is one of the most efficient methods for deriving the factor weights [19,22]. For a rating classification problem, the number of possible attribute sets in rule preconditions totals $2^{|A|}-1$, which is equivalent to the number of times traditional data mining techniques need to train the datasets in order to obtain the weight for each evidence body. This is very time-consuming and computationally prohibitive for real-time recommendation systems. In this paper, we employ a NN-based method, which requires training only once with need-rating data to derive the evidence weights.

For need-rating data *I*, we first find a trained neural network *N* with the entire set of attributes *A*, $A = \{A_1,\ldots, A_h,\ldots, A_{|A|}\}$, as its input. Then the output ratings for $|O|$ customers in data table *I*, denoted as $NC = \{nc_1,\ldots, nc_k,\ldots, nc_{|O|}\}$, $nc_k \in C$, $k = 1,2,\ldots,|O|$, can be calculated. Suppose the real ratings of the customers are $RC = \{rc_1,\ldots, rc_k,\ldots, rc_{|O|}\}$, where $rc_k \in C$ is the rating of the *k*th customer in *O*. The network's accuracy, denoted d_0 can be calculated as follows:

$$d_0 = \sum_{k=1}^{|O|} v_k,\ v_k = \begin{cases} 1 \text{ If } nc_k \text{ is the same as } rc_k \\ 0 \text{ Otherwise} \end{cases}.$$

The weights of evidence bodies can then be computed according to the classification experts in CES. Suppose the precondition of classification expert R_i consists of B_i, $B_i = \{A_{i,1}, A_{i,2}, \ldots A_{i,|B_i|}\}$, the accuracy d_i without the attributes in B_i are computed by simply setting the connection weights from the input attributes $\{A_{i,1}, A_{i,2} \ldots A_{i,|B_i|}\}$of the trained network to zero. Finally, the difference between d_i and d_0 is found to measure the influence of attribute set $\{A_{i,1}, A_{i,2}, \ldots A_{i,|B_i|}\}$ to the classification. The greater the influence of the attribute set is to the classification, the bigger is its weight.

Key steps of the NN algorithm are outlined below.

(1) Let $A = \{A_1,\ldots, A_h,\ldots, A_{|A|}\}$ be the set of all input attributes and RC be the real ratings.
(2) Train the neural network *N* to maximize the network accuracy d_0 with *A* as input such that it achieves a network as accurately as possible.
(3) For $i = 1, 2,\ldots, |\mathrm{CES}|$, let N_i be a network whose weights are as follows:
    (a) For all the inputs except $\{A_{i,1}, A_{i,2}, \ldots A_{i,|B_i|}\}$, assign the weights of N_i equal to the weights of *N*.
    (b) Set the weights from input $\{A_{i,1}, A_{i,2}, \ldots A_{i,|B_i|}\}$to zero.

Compute the output of network N_i, denoted as $\mathrm{NC}_i = \{\mathrm{nc}_{i,1},\ldots, \mathrm{nc}_{i,k},\ldots, \mathrm{nc}_{i,|O|}\}$, and the accuracy of N_i, denoted as d_i.

(4) Compute the influence of B_i to the network accuracy: $w_i = d_0 - d_i$.
(5) If $i \geq |\mathrm{CES}|$, go to (6), otherwise, set $i = i + 1$ and go to (3).
(6) The derived $\{w_1,\ldots, w_i,\ldots, w_{|\mathrm{CES}|}\}$ are the weights of the evidence bodies, where w_i is the weight of evidence body E_i.

The proposed NN method makes it possible to derive the weights quickly and efficiently without falling into the trap of local minimum. It therefore attains more accurate evidence weights and improves the efficiency of the rating classification model.

3.4. Construct rating classifier for potential customers

Given the classification experts, evidence bodies, and evidence weights, the recommendation system is ready to predict the ratings of customers. For a potential customer *c*, the system identifies the necessary classification experts first. For example, R_i and R_j are two matching classification experts whose preconditions are *'Age = Middle'* ∧ *'CPU = Average'* ∧ *'Battery = Good'* for R_i, and *'Age = Middle'* ∧ *'CPU = Average'* for R_j. R_j could be extracted from customers whose preconditions satisfy *'Age = Middle'* ∧ *'CPU = Average'* ∧ *'Battery = Good'* or

'$Age = Middle$' ∧ '$CPU = Average$' ∧ '$Battery \neq Good$', while R_i is extracted from customers whose preconditions satisfy '$Age = Middle$' ∧ '$CPU = Average$' ∧ '$Battery = Good$'. We found R_i to contain more specific information and is more precise to use, thus R_j is not necessary for predicting the ratings of customer c.

To combine evidence bodies, we apply the evidential reasoning method proposed by Yang and colleagues [40]. The evidential reasoning method first transforms the evidence bodies into basic probability masses by combining the normalized evidence weights and the belief degrees. Then, the basic probability masses are combined into an aggregated basic probability assignment. Finally, the aggregated probability assignments are normalized. For customer c, we assume the matching classification expert set consists of S classification experts, $R_c = (R_{c,1}, \ldots, R_{c,u}, \ldots, R_{c,S})$, $R_{c,u} \in \mathrm{CES}$, $u = 1, 2, \ldots, S$. The corresponding evidence bodies and weights are $\mathrm{ES}_c = (E_{c,1}, \ldots, E_{c,u}, \ldots, E_{c,S})$, and $W_c = (w_{c,1}, \ldots, w_{c,u}, \ldots, w_{c,S})$. Rating classification of new customers can be supported by matching its characteristics to one of the classification experts. The matching may lead to one of three situations:

(a) customer c matches one classification expert,
(b) customer c matches more than one classification expert,
(c) customer c does not match any classification experts.

When no classification experts match the new customer, we will not classify the new customer into any class, that is, the proposed model is unable to predict the ratings of customer c. The recommendation system will communicate with the customer and ask for more detailed inputs. In case (a), $S = 1$, there is only one expert $R_{c,1}$ can be used, we employ $R_{c,1}$ to predict the ratings of customer c. In case (b), R_c consists of more than one expert, we use the following procedure to attain classification information.

First, the weights of evidence bodies are normalized to form W'_c, $W'_c = (w'_{c,1}, \ldots, w'_{c,u}, \ldots, w'_{c,S})$, by the following unitary function:

$$w'_{c,u} = \frac{w_{c,u}}{\sum_{l=1}^{S} w_l}, u = 1, 2, \cdots, S$$

Second, calculate the basic probability masses through multiplying W'_c by the evidence bodies in ES_c. Third, calculate the aggregate probability assignment by combining the basic probability masses using the formula proposed by Yang and colleagues [40].

The integrated strategy has two main advantages. First, it ensures that we take all of the essential multi-class information about the new customer into account. The comprehensive utilization of multi-class information plays an important role in constructing an accurate recommendation system. Second, the integrated probability assignment provides the probabilities of possible ratings given by customers after consumption. The ratings together with their respective probabilities allow the recommendation to be more flexible for online stores.

4. Experimental study

4.1. Data

The raw data in our experiment come from online stores which sell Inspiron 1525 laptops. For each customer, we first collect his rating of the laptop, and then retrieve the demographic profile, as well as need and preference statements of the customer. In the process of knowledge discovery from online reviews, the existence of fake reviews is an intractable problem [15,29]. Often, fake reviews are entirely negative or positive. Therefore, to avoid the impact of fake reviews on the accuracy of rating classification, each of the co-authors individually identifies the reviews with complete positive or negative comments, and then collectively decides whether to remove such reviews from the data set. The attributes of age and expertise (computer knowledge) on laptops form customers' demographic profile. Customers' needs and preferences are characterized by CPU speed, battery, audio card, and video card. The attributes and ratings together with possible values are presented in Table 2.

Customers' needs and preferences are extracted by the inverse analysis. For example, if a customer claims that the battery life is insufficient or he/she prefers a laptop with long battery life, we infer that the customer needs a laptop with good battery life. If the customer mentions nothing or claims that the battery life is acceptable, we assume that the laptops with an average battery life will meet his/her needs. Ratings represent customers' evaluation about the product after their usage. From Dell.com and Bestbuy.com we collected 507 need-rating records, of which 405 cases are randomly selected to create the training data, and the remaining 102 records form the testing data. Only the training data is submitted to construct the rating classifier for the laptop. Examples of the need-rating data are shown in Table 3.

Table 2
Attribute description.

Attribute	Profile		Customer need				Rating
	Age	Expertise	CPU	Battery	Audio	Video	
Attribute value	≤24 (Y), 24–34 (M), ≥35 (O)	Average (A), Good (G)	Average, Good	Average, Good	Average, Good	Average, Good	Poor (P), Average, Good

Table 3
Examples of the need-rating data.

Customer	Age	Expertise	CPU	Battery	Audio	Video	Rating
c_1	Y	G	A	G	A	A	G
c_2	O	A	A	A	A	G	P
c_3	M	G	A	A	G	A	A
c_4	M	G	A	A	G	A	G
…	…	…	…	…	…	…	…
c_{505}	O	A	G	A	A	G	G
c_{506}	Y	A	A	A	A	G	P
c_{507}	M	G	A	G	G	A	G

4.2. Mine need-rating rules

To mine classification rules from data, the thresholds of support and confidence degrees are usually set to some small numbers since this can remove the ineffective rules and improve the quality of classifiers. In the rating classification problem, the values of support and confidence thresholds depend on the uncertainty of the need-rating data. This paper sets the two thresholds to 0.05 and 0.1 respectively, which are normally used in most of the associative classification methods [37]. From the need-rating data, we use the proposed method to discover 377 rules. Among all extracted rules, we found 87.3% of them conflict with one another. Twenty-two of the 377 rules are presented in Table 4, which will be used to illustrate the rating classifier construction procedure.

Most of the traditional associative classification methods are designed to find only the rules with the highest confidence level, and use them to classify new objects. However, in our case the conflicting rules are retained through the classification expert pruning method. For example, in Table 4, r_{12}, r_{13}, and r_{14} are three conflicting rules who share the same precondition but different rating results:

'$Expertise = G$' ∧ '$CPU = A$' ∧ '$Battery = G$' → '$Rating = G$';
'$Expertise = G$' ∧ '$CPU = A$' ∧ '$Battery = G$' → '$Rating = P$';
'$Expertise = G$' ∧ '$CPU = A$' ∧ '$Battery = G$' → '$Rating = A$'.

Traditional methods will only select r_{14} as the classification rule due to its high confidence level. In this research we take three phases to derive the multi-class information from these rules.

The first step is to combine the conflicting rules to form classification experts and the corresponding evidence bodies. For example,

Table 4
Rules discovered from the need-rating data.

Rule	Age	Expertise	CPU	Battery	Audio	Video	Rating	Confidence
r_1	*	A	*	*	*	*	G	0.56
r_2	*	A	*	*	*	*	P	0.16
r_3	*	A	*	*	*	*	A	0.28
r_4	Y	A	*	*	*	*	P	0.31
r_5	Y	A	*	*	*	*	A	0.50
r_6	O	*	G	*	*	*	G	0.93
r_7	M	*	A	*	*	*	P	0.31
r_8	M	*	A	*	*	*	A	0.56
r_9	O	A	G	*	*	*	G	0.93
r_{10}	M	G	A	*	*	*	P	0.31
r_{11}	M	G	A	*	*	*	A	0.56
r_{12}	*	G	A	G	*	*	G	0.24
r_{13}	*	G	A	G	*	*	P	0.28
r_{14}	*	G	A	G	*	*	A	0.48
r_{15}	*	G	A	G	*	A	G	0.24
r_{16}	*	G	A	G	*	A	P	0.28
r_{17}	*	G	A	G	*	A	A	0.48
r_{18}	M	*	A	G	*	*	P	0.31
r_{19}	M	*	A	G	*	*	A	0.54
r_{20}	*	*	A	*	A	G	G	0.44
r_{21}	*	*	A	*	A	G	A	0.37
r_{22}	Y	*	*	G	*	*	G	1.00

rules r_{12}, r_{13}, and r_{14} form classification expert R_7. Overall, 174 classification experts are derived from the 377 rules. Table 5 shows 11 of such classification experts, which are based on data from Table 4. In Table 5, Θ represents the frame of discernment for the need-rating data. The numbers in the last four columns represent the confidence (belief) degrees associated with each rating. For example, R_1 implies that the customers with average computer knowledge will rate Inspiron 1525 laptop as *'Poor'* with 16% probability, *'Average'* with 28%, and '*Good*' with 56%. A positive Θ shows the probability that the classification expert cannot predict the responses of customers with such specific precondition. As can be seen in Table 5, there are two kinds of classification experts. One is the classification experts which assign almost all of the belief degrees to one rating. For example, classification expert R_3 implies that customers who are older than 35 and need a laptop with good CPU speed will give a *'Good'* rating with 93% probability for Inspiron 1525 laptop. The other is the classification experts whose ratings vary. For example, the belief degrees of classification experts R_8 are assigned more evenly to all three ratings, *'Poor'*, *'Average'*, and *'Good'*, respectively. The evidence bodies given by such classification experts provide useful information about possible customer ratings for the product.

The second step prunes the inferior classification experts. In Table 5, classification expert R_6 is inferior to R_4 because it gives the same classification information, but contains a more complex precondition. Therefore, R_6 is removed from the set of classification experts. Likewise, classification experts R_5 and R_8 are also removed because they are inferior to R_3 and R_7, respectively.

Table 5
Examples of classification experts.

Classification Expert	Precondition (P)						Evidence Body (E)			
	Age	Expertise	CPU	Battery	Audio	Video	P	A	G	Θ
R_1	*	A	*	*	*	*	0.16	0.28	0.56	0
R_2	Y	A	*	*	*	*	0.31	0.5	0	0.19
R_3	O	*	G	*	*	*	0	0	0.93	0.07
R_4	M	*	A	*	*	*	0.31	0.56	0	0.13
R_5	O	A	G	*	*	*	0	0	0.93	0.07
R_6	M	G	A	*	*	*	0.31	0.56	0	0.13
R_7	*	G	A	G	*	*	0.28	0.48	0.24	0
R_8	*	G	A	G	*	A	0.28	0.48	0.24	0
R_9	M	*	A	G	*	*	0.31	0.54	0	0.15
R_{10}	*	*	A	*	A	G	0	0.37	0.44	0.19
R_{11}	Y	*	*	G	*	*	0	0	1.00	0

Table 6
The final classification expert set.

Classification Expert	Precondition (P)						Evidence body (E)			
	Age	Expertise	CPU	Battery	Audio	Video	P	A	G	Θ
R_2	Y	A	*	*	*	*	0.31	0.5	0	0.19
R_3	O	*	G	*	*	*	0	0	0.93	0.07
R_4	M	*	A	*	*	*	0.31	0.56	0	0.13
R_7	*	G	A	G	*	*	0.28	0.48	0.24	0
R_9	M	*	A	G	*	*	0.31	0.54	0	0.15
R_{10}	*	*	A	*	A	G	0	0.37	0.44	0.19
R_{11}	Y	*	*	G	*	*	0	0	1.00	0

Finally, the unnecessary classification experts are pruned using the database coverage method. The database coverage threshold is set to 4, a number frequently used by other researchers [35]. If a classification expert is necessary to four or more records in the need-rating database, it is retained; otherwise, removed. After all three phases, the original 174 classification experts are reduced to 45 classification experts, which are capable of covering the 405 customers in the training set. In Table 5, the classification expert R_1 is removed by the database coverage method because it is necessary for only two customers' records. As a result, the set given in Table 5 is reduced to Table 6.

4.3. Measure the importance of evidence bodies

After the classification experts are developed from the need-rating data, the corresponding evidence bodies are evaluated using the neural network method. We adopt the radial basis function (RBF) neural network [4] to perform this task. The RBF neural network architecture, which is designed to solve classification problems similar to the radial basis function implemented in the software system MATLAB 7.0, has a single hidden layer with Gaussian function. Using eight data sets from the public machine learning repository [28] to test the efficiency of the RBF neural network for the calculation of attribute weights, we found the average runtime is only 7.87% of the rough set method employed by CSMC [23]. Furthermore, the computational results are comparable between the two methods. The high-speed neural network method significantly improves the capability of the rating classifier. Following the procedures described in Section 3.3, we obtain the weights of all evidence bodies. Table 7 shows the weights of the evidence bodies associated with the classification experts given in Table 6, where E_i is the evidence body given by classification expert R_i, $i \in \{2, 3, 4, 7, 9, 10, 11\}$.

In the need-rating database, 20 more customers will be misclassified if we remove attributes CPU, Audio, and Video from the need-rating data. This implies that the attribute set {CPU, Audio, Video} only has a small classification power, as attested by the small weight, 20, found for evidence body E_{10}. On the contrary, the attribute set {Age, CPU, Battery} can distinguish customers' ratings to a great extent. If we remove the three attributes from the need-rating data, an additional 91 customers will receive the wrong classification. As a result, the weight of the evidence body E_9 is 91, the largest of all. Other weights are derived similarly.

4.4. Construct rating classifier for potential customers

Using the entire 45 classification experts, evidence bodies, and their weights, we are able to construct a comprehensive classifier to predict the ratings of customers with different characteristics. For illustration, we use the classification experts in Table 6 and the

Table 7
The weights of the evidence bodies derived by the neural network.

Evidence body	E_2	E_3	E_4	E_7	E_9	E_{10}	E_{11}
Weight	37	82	82	61	91	20	53

Table 8
A potential customer.

Customer	Age	Expertise	CPU	Battery	Audio	Video
c	*M*	*G*	*A*	*G*	*A*	*G*

weights in Table 7 to predict the ratings of customer c with the characteristics given in Table 8. The matching classification experts, R_4, R_7, R_9, and R_{10}, are found first. Among them, R_4 is not applied because classification expert, R_9 is more specific than R_4.

Thereafter, we combine the multiple evidence bodies given by the three classification experts to calculate the aggregate multi-class classification information. Since the weights for E_7, E_9, and E_{10} are 61, 91, and 20, respectively, which are normalized to .35, .53, and .12 respectively. The evidence bodies used to classify customer c together with their weights are presented in Table 9. Following the method in Section 3.4, we are able to derive the aggregate multi-class classification results as shown in Table 10.

We thus can predict the rating of customer c. Furthermore, the reason why the product is or is not recommended to customer c is lucid. If the product is recommended to customer c, he/she will rate the product as 'Poor' with 26% probability and 'Average' with 53% probability. The chance he/she will consider the product to be "Good" is only 10%. The product thus should not be recommended to the customer.

For the 102 customers in the test data set, two kinds of classification results are obtained by the rating classifier. The first kind of results assigns most of the probabilities to one rating, which makes estimating customer response easy. For example, in Table 11 the predicted ratings for the four customers are 'Good', 'Good', 'Good', and 'Average' with probabilities of 94%, 79%, 70%, and 73%, respectively. The decisions are easy to make, that is, recommend the product to customers c_{412}, c_{435}, and c_{478}, but not to c_{501}.

However, indistinct ratings may also take place, resulting in a less valuable recommendation. For example, the ratings in Table 12 do not provide definite recommendations. Under such circumstances, additional measures may be taken to ensure customer satisfaction. For example, additional information may be elicited to obtain more accurate *needs* information and *preferences* data; a greater discount or a warranty may also be offered to increase the odds of satisfaction.

Table 13 provides accuracy comparisons among different methods. The experiments of the decision tree method (C4.5) and support vector machine (SVM) algorithm were carried out using the *Weka* software system, which is an open source tool for machine learning [39]. The classification based on associations (CBA) algorithm was studied using the software developed by the authors in [26], and the combination strategy for multi-class classification (CSMC) was implemented by software system MATLAB 7.0.

Due to the existence of conflicting ratings, it is hard for traditional methods to mine useful multi-class patterns and construct accurate classifiers. On the contrary, the proposed method can deal with the uncertain environment elegantly. It retains useful conflicting information, integrates conflicting rules to form classification experts, and eventually builds the rating classification model. This explains why the proposed method can attain more accuracy than conventional methods.

Table 9
The set of evidence bodies that matches the target customer.

Evidence body	*P*	*A*	*G*	Θ	Weight
E_7	0.28	0.48	0.24	0	0.35
E_9	0.31	0.54	0	0.15	0.53
E_{10}	0	0.37	0.44	0.19	0.12

Table 10
Rating classification results of the potential customer *c*.

Poor	Average	Good	Θ
0.26	0.53	0.10	0.11

Table 11
Customers whose ratings are easily predicted.

Customer	Age	Expertise	CPU	Battery	Audio	Video	P	A	G	Θ
c_{412}	*Y*	*G*	*A*	*G*	*A*	*A*	0.02	0.04	0.94	0.00
c_{435}	*Y*	*G*	*G*	*G*	*A*	*A*	0.01	0.12	0.79	0.08
c_{478}	*O*	*A*	*G*	*A*	*A*	*G*	0.08	0.14	0.70	0.08
c_{501}	*Y*	*A*	*A*	*A*	*G*	*G*	0.05	0.73	0.01	0.21

Table 12
Customers with inconclusive ratings.

Customer	Age	Expertise	CPU	Battery	Audio	Video	P	A	G	Θ
c_{417}	*M*	*G*	*G*	*A*	*A*	*G*	0.03	0.38	0.27	0.32
c_{452}	*Y*	*A*	*G*	*G*	*A*	*A*	0.00	0.26	0.41	0.33
c_{469}	*Y*	*A*	*G*	*G*	*A*	*G*	0.00	0.39	0.16	0.45
c_{504}	*M*	*G*	*A*	*G*	*G*	*A*	0.25	0.47	0.28	0.00

Table 13
Comparison of classification accuracy.

Method	C4.5	SVM	CBA	CSMC	Proposed model
Accuracy	0.706	0.755	0.686	0.794	0.824

Nomenclature
C4.5: The decision tree method
SVM: Support vector machine algorithm
CBA: Classification based on associations algorithm
CSMC: Combination strategy for multi-class classification.

5. Conclusions and future work

The recommendation system is an important tool to offer personalized service and maximize customer satisfaction. Current literature regards a recommendation system as a success if a potential customer takes the advice and purchases the recommended product. We argue that a truly successful recommendation system should be the one that maximizes the customers' after-sale satisfaction, not one that just lures customers into the act of purchasing. We emphasize that a good recommendation system not only considers what the customer needs, but also ensures customer's contentment. The main contributions of this research are twofold. First, we make a distinction between the customer purchase and the customer endorsement. When a customer follows advice to purchase a product (DO), it does not imply that the person is truly pleased (FEEL) with the decision he/she made. Second, to maximize a customer's satisfaction level, we propose a more effective and efficient rating classification model based on the customer's profile and feedback. The associative classification method proposed in this research is capable of mining multi-class information from the need-rating data. It predicts the appeal of the specific product to the customer through integrated utilization of information, and the recommendation is meticulous and valuable.

Despite the contribution of this research, there are limitations, and further works can be done. The first important work is to investigate the factors that impact a customer's feelings. Many attributes such as the demographic and psychological characteristics, purchase and consumption environment, and customers' expectation, may well have significant influence on customers' feelings toward a specific product. Therefore, it is crucial to indentify the factors important for modeling rating classification, so as to predict the customer's satisfaction level effectively.

Another work is to elicit customers' needs and preferences. The rating classification aims to recommend the right products based on customers' characteristics to achieve high satisfaction levels. Therefore, the validity of customers' needs and preferences has an important implication on the effectiveness of the recommendation system. Oftentimes consumers do not have clear needs and preferences. Therefore, finding an effective way to facilitate customers to express their true needs and preferences is essential for the recommendation systems.

6. Epilogue

Professor W. W. Cooper, a pioneer researcher in management, has made a significant impact on the fields of decision sciences, operational research, accounting, marketing, and human resource management. Among his contributions, Professor Cooper has paid much attention to the research in the area of marketing. He developed innovative models to optimize resource allocation for alternative media advertising [5]. In the 1960s, he and his associates built a strategic decision model, DEMON, for marketing new products [6,7]. His idea of creating a decision support system to aid with marketing decision making inspires our pursuit of this research.

Information technologies, especially Internet technology, bring significant influence to the traditional marketing environment and changes in the direction of research. As early as 1985, Cooper and his colleagues [8] realized the importance of information technology to marketing research. They argued that researchers and practitioners should handle the “problems that may arise for the relations between marketing management and marketing research because of the rapidly increasing use of personal computers.” Indeed, as the Internet becomes a main part of modern society and online shopping develops into a daily activity, online recommendation systems become ubiquitous and widely utilized by practitioners to improve their revenues. Our research focuses directly on the improvement of recommendation systems.

In investigating the rating classification problem, we follow Dr. Cooper's insights about marketing research. In his opinion, when dealing with decision-making problems under uncertainty, the marketing model should be “simple and intuitive, and easy to understand by both academic researchers and practitioners.” Our research proposed a novel associative classification model to handle the rating classification problem. The proposed model is easy to understand, capable of dealing with uncertainty, and more practical and logical than existing techniques. Therefore, the associative classification model can be understood and used by practitioners straightforwardly. Moreover, the outcome of our research is not limited to only the classification results. According to Dr. Cooper, “simply predicting what will happen in the future is of less interest to managers than knowing what has to be changed, and by how much, to achieve their goals.” This paper follows Professor Cooper's guideline by detecting the probabilities of customers' satisfaction levels beforehand. Such an approach gives the basis for marketers to adopt various marketing strategies to achieve high satisfaction levels. We attribute our recommendation system, with the ultimate goal of marketing online products to maximize customer satisfaction, to Dr. Cooper's pioneering thinking.

Acknowledgements

The authors thank the editors and two anonymous reviewers for their insightful comments. This work was supported by the National Science Foundation of China (Project No.70672097) and the State Key Program of National Natural Science of China (Project No.70631003).

References

[1] R. Agrawal, R. Srikant, Fast algorithms for mining association rules in large databases, 20th International Conference on Very Large Data Bases, Santiago, 1994, 1994.
[2] A. Ansari, S. Essegaier, R. Kohli, Internet recommendation systems, Journal of Marketing Research (JMR 37 (3) (2000) 363–375.
[3] A.V. Bodapati, Recommendation systems with purchase data, Journal of Marketing Research (JMR 45 (1) (2008) 77–93.
[4] D.S. Broomhead, D. Lowe, Multivariable functional interpolation and adaptive networks, Complex Systems 2 (1988) 321–355.
[5] A. Charnes, W.W. Cooper, A constrained game formulation of advertising strategies, Econometrica 22 (3) (1954) 511–512.
[6] A. Charnes, W.W. Cooper, J.K. DeVoe, D.B. Learner, Demon: decision mapping via optimum go-no networks-a model for marketing new products, Management Science 12 (11) (1966) 865–887.
[7] A. Charnes, W.W. Cooper, J.K. DeVoe, D.B. Learner, DEMON, Mark II: an extremal equations approach to new product marketing, Management Science 14 (9) (1968) 513–524.
[8] A. Charnes, W.W. Cooper, D.B. Learner, Management science and marketing management, Journal of marketing research 49 (1985) 93–105.
[9] Y. Chen, D. Liginlal, A maximum entropy approach to feature selection in knowledge-based authentication, Decision Support Systems 46 (1) (2008) 388–398.
[10] K.W. Cheung, J.T. Kwok, M.H. Law, K.C. Tsui, Mining customer product ratings for personalized marketing, Decision Support Systems 35 (2) (2003) 231–243.
[11] J.A. Chevalier, D. Mayzlin, The effect of word of mouth on sales: online book reviews, Journal of Marketing Research 43 (3) (2006) 345–354.
[12] Y.H. Cho, J.K. Kim, S.H. Kim, A personalized recommender system based on web usage mining and decision tree induction, Expert Systems with Applications 23 (3) (2002) 329–342.
[13] E.K. Clemons, G. Gao, L.M. Hitt, When online reviews meet hyper differentiation: a study of craft beer industry, Journal of Management Information 23 (2) (2006) 149–171.
[14] W.W. Cooper, L.M. Seiford, K. Tone, Data envelopment analysis: a comprehensive text with models, applications, references and DEA-solver software, Kluwer Academic Publishers, Boston, 2000.
[15] C. Dellarocas, The digitization of word of mouth: promise and challenges of online feedback mechanisms, Management Science 49 (10) (2003) 1407–1424.
[16] M. Deshpande, G. Karypis, Item-based top-N recommendation algorithms, ACM Transactions on Information Systems 22 (1) (2004) 143–177.
[17] D. Goldberg, D. Nichols, B.M. Oki, D. Terry, Using collaborative filtering to weave an information tapestry, Communications of the ACM 35 (12) (1992) 61–70.
[18] J.L. Herlocker, J.A. Konstan, J. Loren, G. Terveen, T. Riedl, Collaborative filtering recommender systems, ACM Transactions on Information Systems 22 (1) (2004) 5–53.
[19] M.Y. Hu, M. Shanker, G.P. Zhang, M.S. Hung, Modeling consumer situational choice of long distance communication with neural networks, Decision Support Systems 44 (4) (2008) 899–908.
[20] Y. Huang, L. Bian, A Bayesian network and analytic hierarchy process based personalized recommendations for tourist attractions over the Internet, Expert Systems with Applications 36 (1) (2009) 933–943.
[21] L.P. Hung, A personalized recommendation system based on product taxonomy for one-to-one marketing online, Expert Systems with Applications 29 (2) (2005) 383–392.
[22] Y. Kim, W.N. Street, An intelligent system for customer targeting: a data mining approach, Decision Support Systems 37 (2) (2004) 215–228.
[23] J. Konstan, B. Miller, D. Maltz, J. Herlocker, L. Gordon, J. Riedl, GroupLens: applying collaborative filtering to usenet news, Communications of the ACM 40 (3) (1997) 77–87.
[24] W.M. Li, J.W. Han, J. Pei, CMAR: Accurate and Efficient Classification Based on Multiple Class-Association Rules, Proceedings of the 2001 IEEE International Conference on Data Mining. 2001: California, 2001.
[25] T.P. Liang, Y.F. Yang, D.N. Chen, Y.C. Ku, A semantic-expansion approach to personalized knowledge recommendation, Decision Support Systems 45 (3) (2008) 401–412.
[26] B. Liu, W. Hsu, Y. Ma, Integrating classification and association rule mining, Proceedings of the Fourth International Conference on Knowledge Discovery and Data Mining (KDD-98), New York, 1998, 1998.
[27] Y.Z. Liu, Y.C. Jiang, X. Liu, S.L. Yang, CSMC: a combination strategy for multi-class classification based on multiple association rules, Knowledge-Based Systems 21 (8) (2008) 786–793.
[28] P.M. Murphy, D.W. Aha, UCI Repository machine learning databases 1996, University of California, Irvine, CA, 1996.
[29] D.-H. Park, J. Lee, eWOM overload and its effect on consumer behavioral intention depending on consumer involvement, Electronic Commerce Research and Applications 7 (4) (2008) 386–398.
[30] J.R. Quinlan, C4.5: programs for machine learning, Morgan: Morgan Kaufmann Publishers, 1993.
[31] R. Rak, L. Kurgan, M. Reformat, A tree-projection-based algorithm for multi-label recurrent-item associative-classification rule generation, Data & Knowledge Engineering 64 (1) (2008) 171–197.
[32] S. Senecal, J. Nantel, The influence of online product recommendations on consumers' online choices, Journal of Retailing 80 (2) (2004) 159–169.
[33] M.-D. Shieh, C.-C. Yang, Multiclass SVM-RFE for product form feature selection, Expert Systems with Applications 35 (1–2) (2008) 531–541.
[34] T. Sueyoshi, G.R. Tadiparthi, Agent-Based Approach to Handle Business Complexity in U.S. Wholesale Power Trading, IEEE Transactions on Power System 22 (2) (2007) 532–542.
[35] F. Thabtah, A review of associative classification mining, The Knowledge Engineering Review 22 (1) (2007) 37–65.
[36] F.A. Thabtah, P.I. Cowling, A greedy classification algorithm based on association rule, Applied Soft Computing 7 (3) (2007) 1102–1111.
[37] F. Thabtah, P. Cowling, S. Hammoud, Improving rule sorting, predictive accuracy and training time in associative classification, Expert Systems with Applications 31 (2006) 414–426.
[38] F.H. Wang, H.M. Shao, Effective personalized recommendation based on time-framed navigation clustering and association mining, Expert Systems with Applications 27 (3) (2004) 365–377.

[39] I.H. Witten, E. Frank, in: M. Kaufmann (Ed.), Data Mining: Practical machine learning tools and techniques, 2005, San Francisco.
[40] J.B. Yang, Y.M. Wang, D.L. Xu, et al., The evidential reasoning approach for MADA under both probabilistic and fuzzy uncertainties, European Journal of Operational Research 171 (1) (2006) 309–343.
[41] A. Zenebe, A.F. Norcio, Representation, similarity measures and aggregation methods using fuzzy sets for content-based recommender systems, Fuzzy Sets and Systems 160 (1) (2009) 76–94.

Yuanchun Jiang received his bachelor's degree in management science and engineering from Hefei University of Technology, Hefei, China. He is a PhD student in the Institute of Electronic Commerce in the School of Management at Hefei University of Technology. He is currently a visiting PhD student in the Joseph M. Katz Graduate School of Business at the University of Pittsburgh. His research interests include decision science, electronic commerce, and data mining. He has published papers in journals such as *Knowledge-Based Systems, Systems Engineering-Theory & Practice*, and *Journal of Systems Engineering.*

Jennifer Shang received her PhD in Operations Management from the University of Texas at Austin. She teaches operations management, simulation, statistics, and process and quality improvement courses. Her main research interests include multi-criteria decision making and its application to the design, planning, scheduling, control, and evaluation of production and service operational systems. She has published in various journals, including *Management Science, Journal of Marketing, European Journal of Operational Research, Decision Support Systems, IEEE Transactions on Engineering Management,* and *International Journal of Production Research*, among others. She has won the EMBA Distinguished Teaching Award and several Excellence-in-Teaching Awards from the MBA/EMBA programs at Katz Business School.

Yezheng Liu is a professor of Electronic Commerce in Hefei University of Technology. Dr. Liu received his PhD in management science and engineering from Hefei University of Technology. He teaches electronic commerce, decision sciences, and information systems. His main research interests include data mining and its application in electronic commerce, decision support systems, and optimization models.