Appl Intell (2008) 29: 279–288
DOI 10.1007/s10489-007-0094-7

A personalized counseling system using case-based reasoning with neural symbolic feature weighting (CANSY)

Sungho Ha

Published online: 6 October 2007

Abstract In this article, we introduce a personalized counseling system based on context mining. As a technique for context mining, we have developed an algorithm called CANSY. It adopts trained neural networks for feature weighting and a value difference metric in order to measure distances between all possible values of symbolic features. CANSY plays a core role in classifying and presenting most similar cases from a case base. Experimental results show that CANSY along with a rule base can provide personalized information with a relatively high level of accuracy, and it is capable of recommending appropriate products or services.

Keywords Personalization · Data mining · Machine learning · Case-based reasoning · Feature weighting · Value difference metric

1 Introduction

In the highly competitive world of rapidly developing technology, companies that merely respond to the basic needs of customers cannot ultimately survive. Therefore, in recent years there has been a growing interest in personalized services from the perspective of one-to-one marketing.

Personalization, a special form of differentiation, when applied to market segmentation, can transform a standard product or service into a specialized solution for an individual. Through personalization, businesses can get to know customer-buying behavior and accordingly, they can develop more appropriate marketing strategies which can attract specific types of customers. As a result, businesses are able to deliver suitable information and products or services more efficiently. Customer satisfaction and loyalty can thus, be enhanced and an increase in customer patronage can create more transaction opportunities and benefits [1].

Data mining techniques, which support personalization, include the need to cluster, classify, and search for data within a very large data space. Among the techniques, neural networks (NN) and case-based reasoning (CBR) can be directly applied to classification purposes without the need of additional transform mechanisms. NN and CBR are able to 'learn' the dynamic behavior of a system over a period of time [2].

Unfortunately, they have disadvantages. The knowledge representation of NN is unreadable and its 'black box' property restricts it from being applied to areas which need proper explanations as well as precise classification. CBR suffers from feature weighting; when it measures the distance between cases, some features should be weighted differently. Many feature-weighted variants of the k-nearest neighbor (k-NN) have been proposed to assign higher weights to more relevant features for case retrieval purposes [3]. Though those variants have been reported as improving their retrieval accuracy regarding some tasks [4], few have been used in conjunction with neural networks.

This was until the proposal of MANN (Memory and Neural Network based learning) [5]. In a hybrid approach of NN and CBR, MANN can calculate the set of feature weights from the trained neural network, which was used to obtain the most similar examples from a case base. MANN supplies the more important features with larger weight values, which improves the retrieval accuracy of CBR. MANN, however, works best in domains in which all features have

S. Ha (✉)
1370 Sangyeok-dong, Buku-gu, 702-701 Daegu, Republic of Korea
e-mail: hsh@mail.knu.ac.kr

numeric values. This is because a normalized Euclidean distance is used to compare examples.

When features have symbolic or unordered values, a more sophisticated treatment of the feature space is required [6, 7]. Therefore, as an algorithm for context mining, we have devised a hybrid system called CANSY (Case-based reasoning and Neural network for SYmbolic features). In CANSY, we adopt trained neural networks for weighting symbolic features and for guiding CBR. This is done by providing global weights to the learning process. We adopt a value difference metric (VDM) in order to measure statistical distances between all possible values of symbolic features.

In this article, we apply CANSY to the building of a personalized counseling system for the cosmetic industry. In the cosmetic industry, existing services depend on a counselor's ability to interact with customers and to provide accurate skin test results. Skilled counselors provide customers with quality service, but unskilled counselors cannot. Companies want to provide customers with stable and high-quality service [8]. To meet these needs, the developed system deploys a case base using a context mining technique (the CANSY system) and an add-on rule base that stores counselor knowledge using data mining techniques. Through this approach, the system can provide personalized information containing appropriate lifestyle, makeup, and beauty products for each customer. The experimental results show that it is possible to achieve a relatively high level of classification accuracy and to maintain such standards without much effort.

The article is organized as follows: In Sect. 2, we introduce the concept of context mining and review various approaches in providing personalized information. Section 3 describes the framework and methodology of the personalized counseling system. Then, we present a Web-based application which has been developed for a cosmetic company in Sect. 4. To validate our system, experimental results are also presented in this section. Finally, we conclude this article by briefly summarizing the study and the direction of future research.

2 Literature review

2.1 Personalization and context mining

Context is a powerful concept in human-computer interaction. It is a set of implicit or explicit stimuli surrounding a specific individual, and it consists of a physical and a social environment that may affect behavior. More specifically, user context is any information, including a user profile or location that can be used to characterize a user [9].

Context information related to personal activation in commercial transaction is classified as time, identity, location, and entity. Time context is usable when a user enters a service zone and if a service is available at that time, the service works for the benefit of the user. Identity context indicates with whom a user is communicating. Location context represents the place of an identity in which a user is interested. Entity context is information that a user may be currently exploiting [10, 11].

Figure 1 shows a personalization framework, based on context mining. The realm of customer information handled by enterprises expands dramatically under an Electronic commerce environment. To achieve electronic customer relationship management, all customer information is stored in a lightly- or heavily-summarized data format in a context data warehouse: It includes customer demographics, preferences, purchase (sales) history, product information, Web navigation patterns, survey responses, and other contextual data.

To utilize contextual information for personalization services, a context database has the following content: entity, entity history, current situation, individual location, and environmental information. Entity is composed of customer profile information. Entity history is used to recommend goods or services through the analysis of customer preferences. Current situation indicates current activity information of a customer, and it is used to recognize the current context. Environmental information is provided to obtain the information of an entity, for example, items selected by the entity.

Contextual data can be analyzed with traditional data mining techniques [12], but new mining techniques (or algorithms) are needed. New mining techniques must handle contextual data which are comprised of nominal or symbolic values. This new mining technique is called *context mining*.

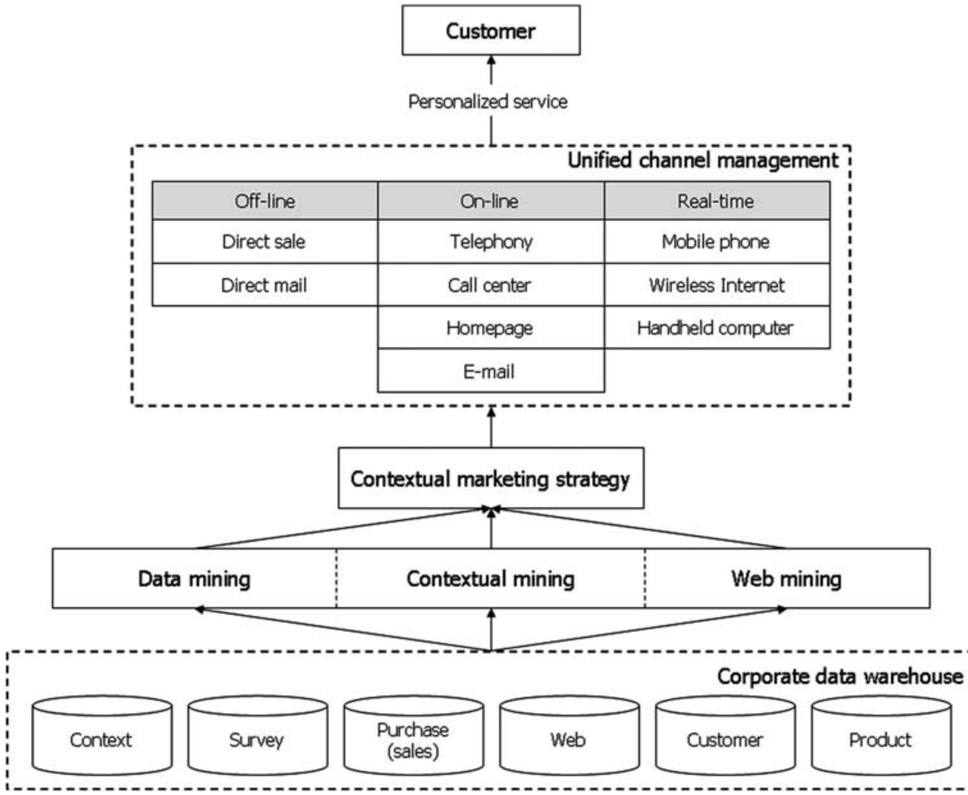
With the analysis results obtained by context mining, Web mining, and data mining, context marketing strategies can be developed and delivered to proper customers. Personalization, as a tool of providing specific information to each customer, is applied to all customer contact channels. Channels are unified in an Electronic commerce environment and management can contact customers any time and any place through off-line, on-line, and real-time channels.

2.2 Personalization techniques

Most personalization techniques aim to personalize Web sites in order to provide personalized services. While a user is visiting and browsing a company's Web site, personalization can provide recommendations based on what products the user examines. Furthermore, personalization can make recommendations based on customers entering words or phrases to describe the type of products they are searching for [13].

Personalization techniques are divided into four major categories: decision rule-based, content-based, collaborative, and learning-agent personalization [14, 15]. Decision

Fig. 1 A personalized service framework based on context mining



rule-based personalization surveys users and then, manually specifies decision rules based on the user demographics or static profiles. It delivers the appropriate content to a particular user based on these rules.

Considerable efforts, however, have been made towards these two directions, such as content-based and collaborative, and their hybrid [16, 17]. Content-based personalization recommends goods or services similar to the customer's previous choices. Personalization gathers information about customer's preferences via the past purchasing history which is stored in databases. Collaborative personalization provides recommendations on the basis of the purchasing preferences of other customers with similar interests or demographics. It is the most successful personalization technology to date [18–21].

There are many commercial and non-commercial sites that implement collaborative personalization systems [22]. For example, Amazon, AlexLit, and StoryCode provide book recommendations; Findory is a personalized news website; Pandora, Last.fm, and Musicmatch are music recommendation systems; Netflix and Hollywood Video provide movie recommendations; and Loomia, Net Perceptions, and Touchstone provide recommendation services for E-commerce.

Learning-agent personalization incorporates Web usage data to discover navigation patterns and to predict user behavior so as to create dynamic user profiles and maintain up-to-date personal preferences.

3 Personalized counseling system

Cosmetic companies own information about their customers, beauty products, and makeup skills based on skin type and skin sensitivity. By using this information, they want to provide customers with stable and high-quality counseling. By the request of the cosmetic company, we have developed a personalization system which has three functions: CANSY (feature-weighted CBR), skin analysis, and personalized counseling. Figure 2 illustrates the framework of the personalized counseling system.

First, we devise a series of questionnaires to collect information regarding customer characteristics (e.g., physical constitution), data concerning personal preferences, life style, and makeup. Then, we build an integrated database which houses answers to questionnaires. Since answers to these questionnaires are not numeric but symbolic, a context mining technique, CANSY, is adopted.

CANSY analyzes customer information by using a series of answers to questionnaires and it evaluates the five basic elements of the Chinese cosmogony (metal, wood, earth, water, and fire) of each customer. In order to determine the skin type, moisture level, and sensitivity level of each customer, information is acquired from questionnaires, physical skin tests, and a rule base which stores specialized knowledge discovered from counselors via data mining techniques. By utilizing the five elements, skin type, and

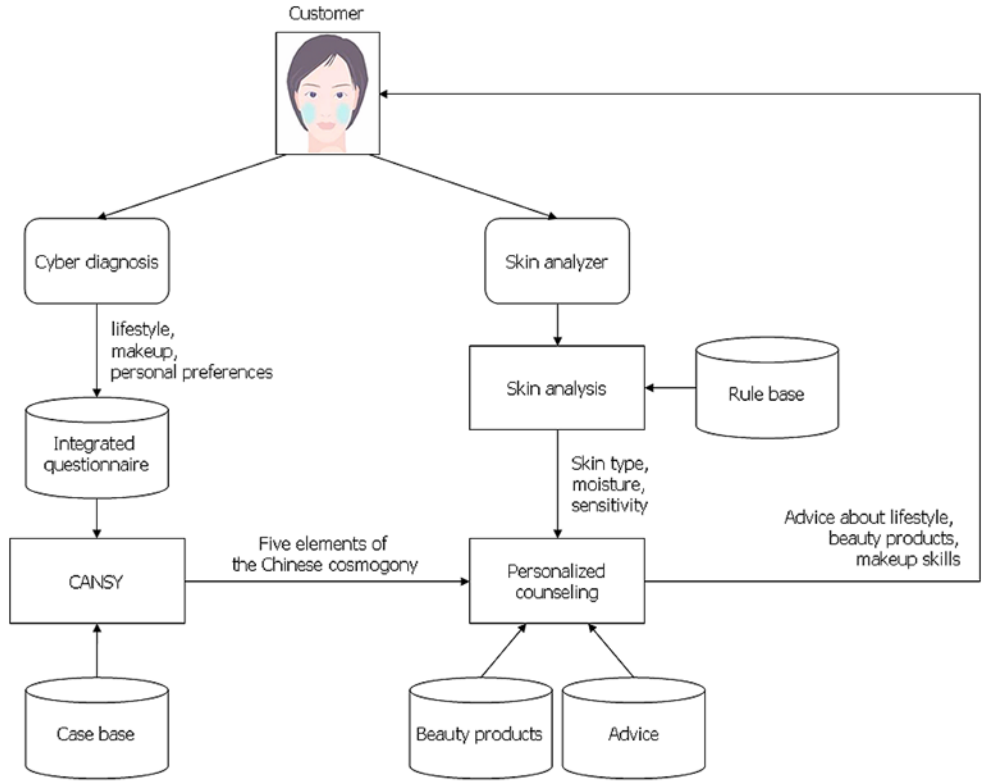


Fig. 2 A framework of a personalized counseling system

skin sensitivity, the counseling system can provide personalized advice regarding makeup and cosmetics that are suitable for each customer.

3.1 Evaluation of the five basic elements by CANSY

In oriental medicine, everything in nature can be classified into one of five basic elements: metal, wood, earth, water, and fire. These five elements are not just the materials to which the names refer, but they are metaphors or symbols that describe how things interact and relate to each other. The five elements contain one's physical constitution and characteristics. If beauty consultants know a customer's type regarding the five elements of the Chinese cosmogony, they can provide personalized services based on his or her own physical constitution.

In this article, CANSY is utilized to cluster and classify symbolic contextual data, whereby all input and target features are symbolic (refer to Fig. 3).

CANSY consists of three phases. The first phase discovers a set of feature weights, which can be acquired by the training of the instances (cases) in a case base. Since symbolic feature values have no distance between values, they cannot be used as the input data of a neural network and the normalized Euclidean distance can not be used as the distance function. Therefore, a transformation mechanism is required in order to use symbolic feature values as input data of a neural network. Before the neural network is 'trained', all symbolic features are transformed into binary-format features.

When the training of a neural network is finished, a set of feature weights from the trained neural network can be obtained. At this time, we use four feature-weighting methods: Sensitivity, Activity, Saliency, and Relevance [3, 23, 24]. Each of these methods calculates the degree of each feature's importance by using the connection weights and activation patterns of the nodes in the trained neural network. The symbolic feature-weighting algorithms are briefly described as follows:

- Input node's sensitivity is calculated by removing the input node from the trained neural network. The sensitivity of an input node is the difference in error between the removal of the feature and when it is left in place.
- Node activity is measured by the variance of the level of activation for the training data. When the activity value of a node varies significantly according to its input value, the activity of the node is high.
- Saliency is measured by estimating the second derivative of the error with respect to weight. Saliency is used to prune neural networks iteratively: that is, to train to reasonable error levels, compute saliencies, delete low saliency weights, and resume training.
- The variance of weight in a node is a good predictor of the node's Relevance. This relevance is a good predictor of the expected increase in error when the node's largest weight is deleted.

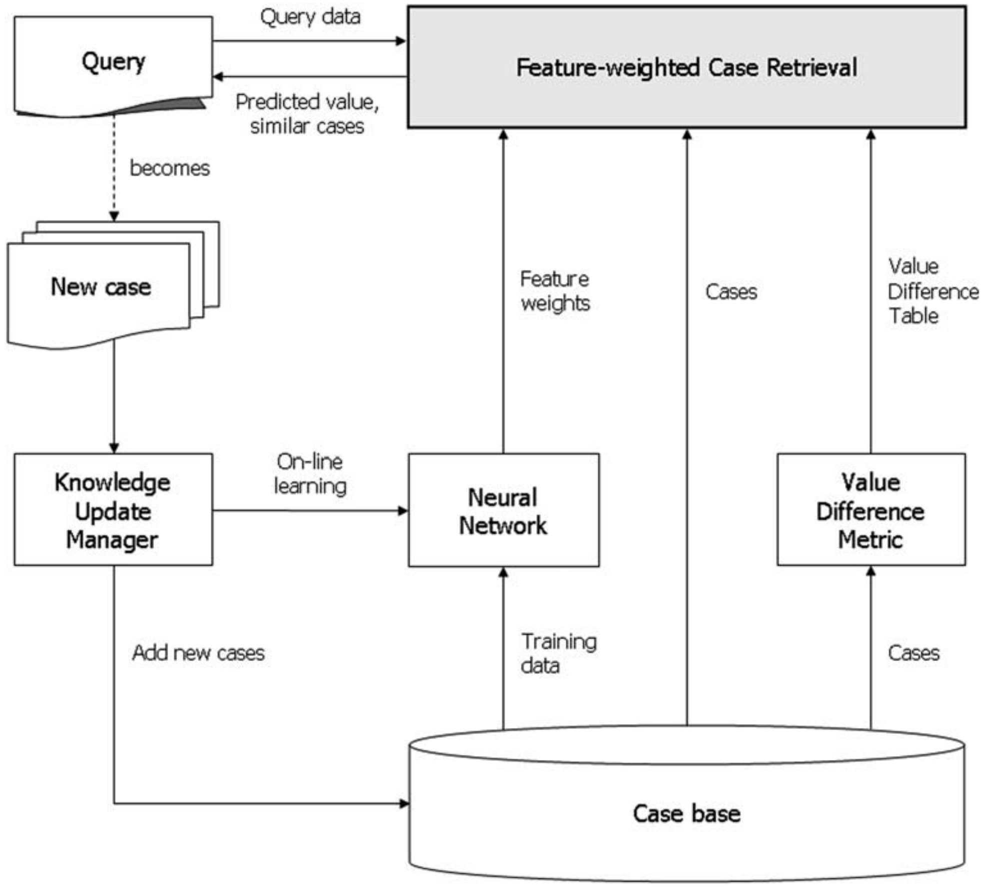


Fig. 3 The structure of CANSY

Table 1 Four feature-weighting algorithms

Algorithm	Input node (x_i)	Binary input node (b_l)	Hidden node (z_j)
Sensitivity	$Se_i = \frac{(\sum_L \lvert P^o - P^i \rvert / P^o)}{n}$		
Activity	$A_i = \sum_{b_l \in x_i} \frac{A_l}{n_i}$	$A_l = \sum_{j=1}^{H} w_{jl}^2 A_j$	$A_j = \sum_{k=1}^{O} w_{kj}^2 \operatorname{var}[g(\sum_{l=0}^{L} w_{jl} b_l)]$
Saliency	$Sa_i = \sum_{b_l \in x_i} \frac{Sa_l}{n_i}$	$Sa_l = \sum_{j=1}^{H} \sum_{k=1}^{O} w_{kj}^2 w_{jl}^2$	
Relevance	$R_i = \sum_{b_l \in x_i} \frac{R_l}{n_i}$	$R_l = w_{jl} R_j$	$R_j = \sum_{k=1}^{O} w_{kj}^2 \operatorname{var}(w_{jl})$

Table 1 summarizes the four feature-weighting algorithms. Notice that L is the set of training data; n is the number of training data; x_i represents original input nodes ($i = 1, \ldots, I$); b_l denotes transformed binary input nodes ($l = 1, \ldots, L$); z_j indicates hidden nodes ($j = 1, \ldots, H$); and y_k signifies output nodes ($k = 1, \ldots, O$). The number of output nodes denotes the number of possible values of the target feature. Each output node represents one possible value of the target feature, namely one of the target classes. n_i is the number of values of the x_i input feature; P^o is the normal value for each training instance classified by a trained neural network and P^i is the modified value classified when an input i is removed; and var() is the variance function.

The second phase is the construction of the value difference table (VDT) from the instances in the case base. This is according to the VDM that defines the distances between the different values of a given feature. For each feature, the VDM is derived statistically based on the instances from the case base, according to (1).

$$\delta(V_1, V_2) = \sum_{i=1}^{n} \left| \frac{C_{1i}}{C_1} - \frac{C_{2i}}{C_2} \right| \tag{1}$$

where V_1 and V_2 are two possible values for a feature and n is the number of classes. C_1 is the total number of times V_1 occurred and C_{1i} is the number of times V_1 was classified into class i. The term C_{1i}/C_1 is the likelihood that an instance will be classified as class i given that its ith feature has value V_1. Thus, (1) computes the overall similarity between two values by finding the sum of the difference of the likelihood for all classifications.

The third phase is case-based reasoning which uses feature weights that are extracted from the trained neural net-

Table 2 A set of features for skin analysis

Q1	Q2	Q3	...	Skin test 1	Skin test 2	Skin test 3	Skin type	Sensitivity
4	2	3	...	56	36	65	4	3
1	4	3	...	34	45	21	1	4

work and the VDT that is constructed by the VDM. The hybrid system of case-based reasoning and neural network improves the retrieval accuracy of CBR. When a new query comes in, the distance, $\Delta(q,x)$, between the query and the case is calculated as follows:

$$\Delta(q,x)=\sum_{i=1}^{N} w_i\delta(q_i,x_i)^r \qquad (2)$$

where q is the query and x is a case from the case base, q_i and x_i are the ith feature values of q and x. w_i is the weight of the ith input feature. $\delta(q_i,x_i)$ is the distance between two values q_i and x_i of the ith input feature. r is usually set to either one or two according to the case base. In this article, we set r to one for all experiments. This ensures that most similar cases to the query are retrieved.

A new query may become a new case. The Knowledge Update Manager (KUM) adds new cases, and it provides them with on-line learning. The KUM trains the neural network to reflect these new cases.

3.2 Evaluating skin type and sensitivity by using a rule base

If beauty consultants know a customer's skin type and sensitivity, they can provide personalized counseling services based on his or her skin conditions. In doing so, a skin test should be conducted. This skin test includes measuring the hydration level of the cheeks, the sebum level of the forehead, and the level of wrinkles of the face. Then, beauticians add a class value which represents skin type and sensitivity. Skin type is classified into one of four categories and sensitivity is classified into one of three groups.

Table 2 shows the features that are used to extract decision rules. A series of questions (Q1, Q2, Q3, and ...), skin type, and skin sensitivity have symbolic values. Several skin test results (# 1, 2, and 3) have numerical values.

We generate decision tree classifiers in order to classify skin type and sensitivity. The resulting decision trees can be transformed into symbolic production rules. These rules reflect the specialized knowledge of beauty counselors, and they are used to determine a customer's skin type and sensitivity level.

3.3 Providing skin-care advice and beauty items

Personalized counseling offers customers skin-care advice regarding lifestyle, makeup, and cosmetics. To provide advice regarding a customer's lifestyle, the counseling system refers to the Advice database with the derived five basic elements. For example, if a customer has a wood value, the personalized counseling system displays wood style advice about leisure, fashion, the interior design of a house, food, and health; Wood-style people are known to be self-confident and upright moral character, so they can enjoy a productive life if they improve their generosity and creativity.

In addition, the system provides advice about makeup based on skin analyses, including skin type and sensitivity. Appropriate beauty products can be recommended based on the five elements, skin type, and sensitivity. Moreover, the system can display information from customers that share similar characteristics. For example, if a customer belongs to the wood type and has dry skin, the system recommends daily and weekly makeup sessions and cosmetics that are geared toward replenishing moisture.

4 Application and evaluation

We developed a personalized counseling system, which is suitable for the web environment. An environment for application development has deployed IBM Websphere as a web application server and Oracle 8.x as a database management system. HTML, Java Script, JavaServer Pages, and Java Beans were chosen as development languages. To interface between web browsers and skin analyzers (skin-tester machines), we developed ActiveX interface modules. The system was designed to supply customers with personalized counseling immediately when they answer questions via the Internet. It can also provide personalized counseling to customers when they undergo a skin test by using a skin analyzer (skin measuring instrument) in a cosmetic shop.

Figures 4 and 5 contain sample screens which are developed for the personalized counseling system. Figure 4 illustrates four screens, which ask customers several questions. For this system, beauty consultants help prepare around 30 questions regarding a customer's lifestyle, makeup preference, and other personal inclinations. Figure 5 shows all advice content regarding leisure, fashion, interior, and food issues that customers will receive, which are based on the Chinese five basic elements.

As a customer's lifestyle changes, new cases continuously accumulate in the case base and customers can receive up-to-date information regarding makeup and beauty

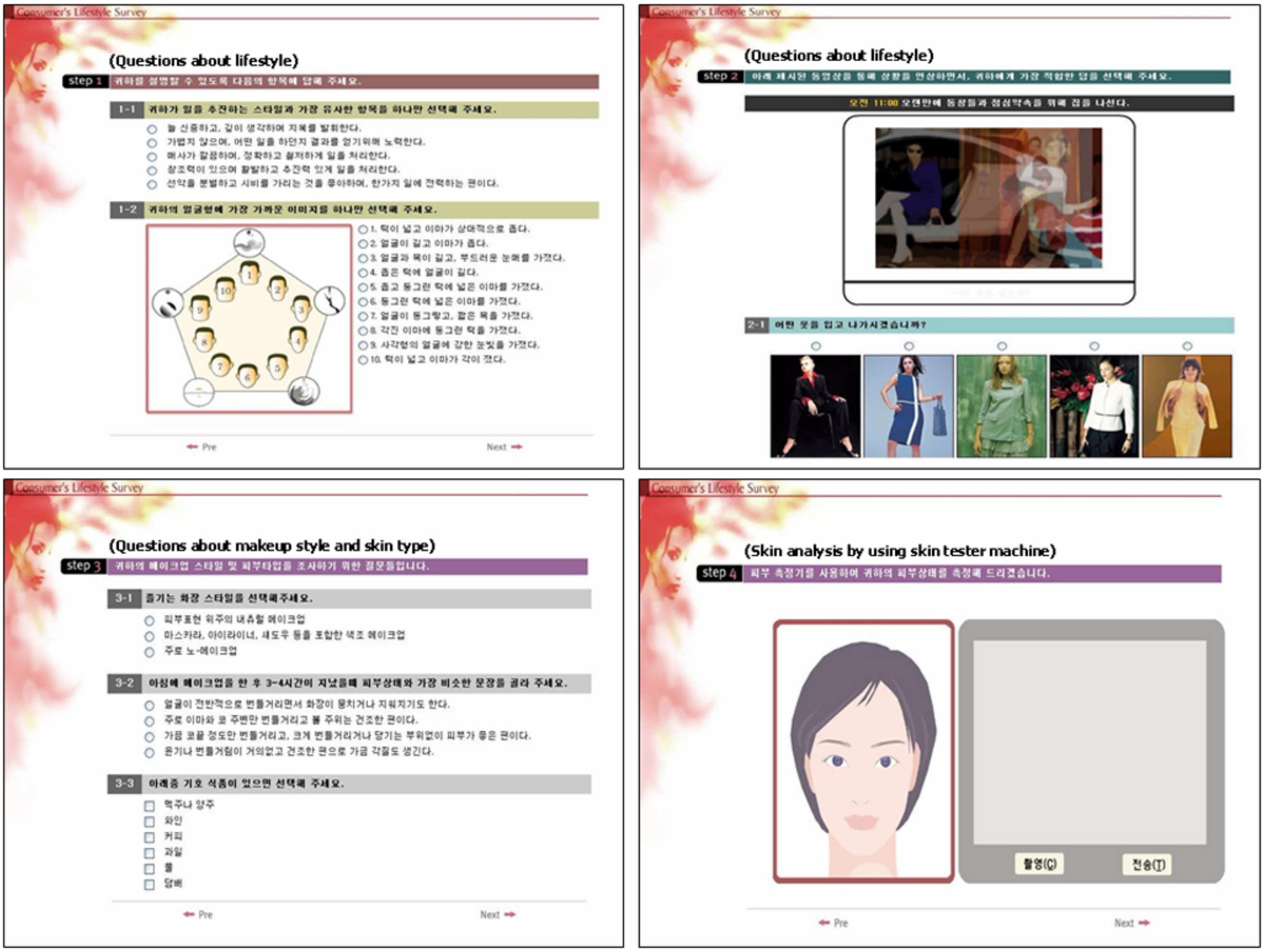


Fig. 4 Customer questionnaires asking personal lifestyle and makeup preferences (Some Korean words were translated into English)

products. This guarantees the continuous satisfaction of customers.

To validate our system, we compared classification accuracy between the uniform k-NN algorithm without feature weighting and feature-weighted k-NN algorithms. In this experiment, beauty consultants collected customer survey data and added five element class values to the set. The dataset had 400 instances, which was divided into a training dataset and a test dataset of 300 and 100, respectively. Experiments were repeated ten times for each k, the number of nearest neighbors.

We created a neural network with one input layer (12 input nodes), one hidden layer (20 hidden nodes), and one output layer (5 output nodes). To train the neural network, we applied the gradient descent algorithm with a momentum and an adaptive learning rate, which was implemented in MATLAB 6 as a widespread training algorithm. In calculating weight values of the input features, the four methods, including Sensitivity, Activity, Salience, and Relevance, were used. To calculate the distance between two symbolic features, the VDT, which was calculated from the VDM, was obtained.

Figures 6 and 7 show the performances of all feature-weighting methods according to varying k, which are odd numbers between 1 and 15. A small neighborhood (k) has better accuracy, since the case base in use is well-organized. In this case, when k increases, the classification accuracy may decrease, since there are many possibilities that distant cases may be included.

Figure 6 illustrates the accuracy level of the feature-weighting methods without the online learning of the KUM. When k is less than five, the four weighting methods show better classification accuracy than the uniform method. When k is larger than five, however, the uniform method outperforms some weighting methods. There are small differences in the classification accuracy between the four feature-weighting methods. As k increases, the classification accuracy fluctuates and decreases. To some extent this problem can be resolved by the dynamic learning feature of the KUM.

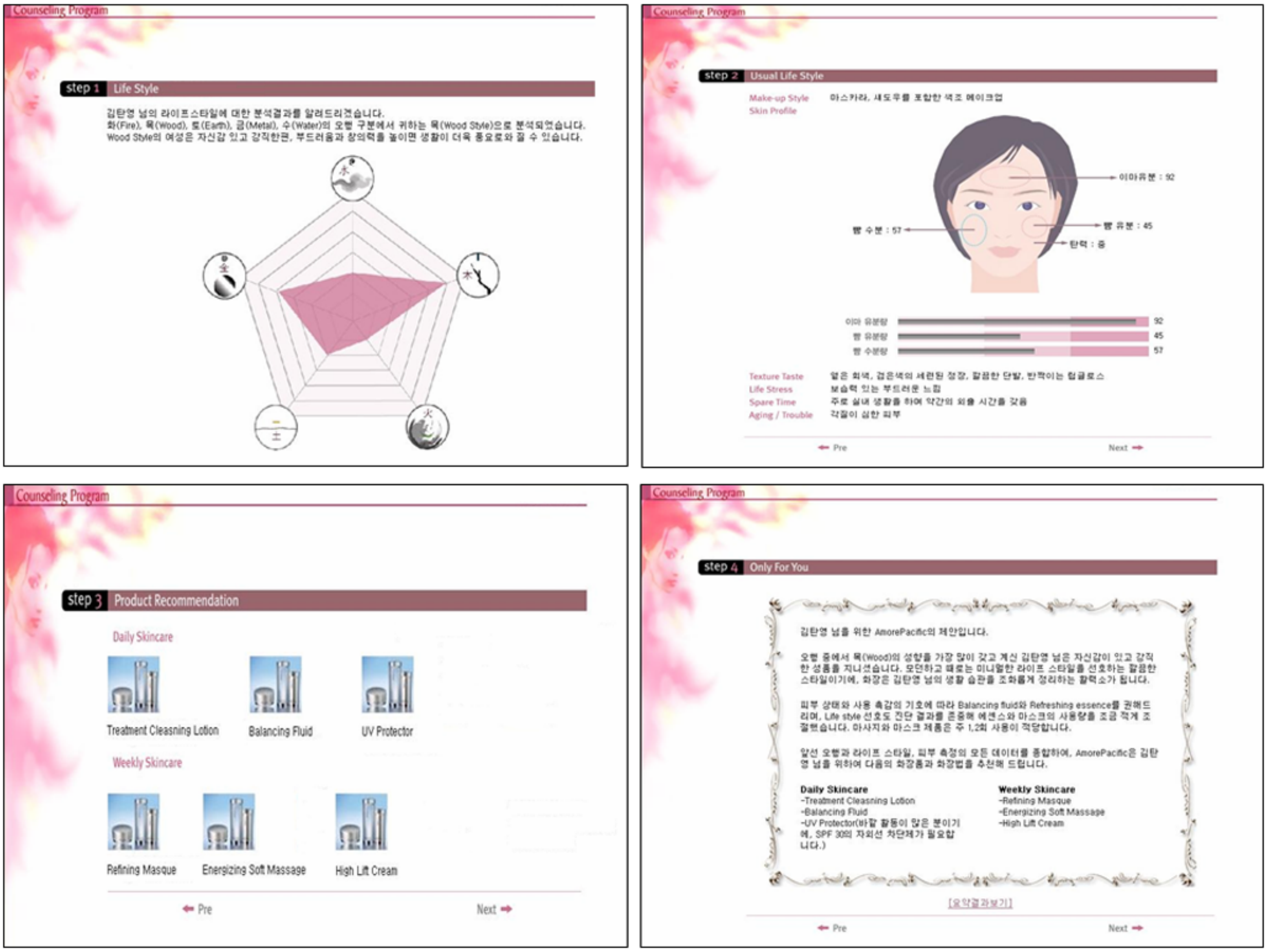
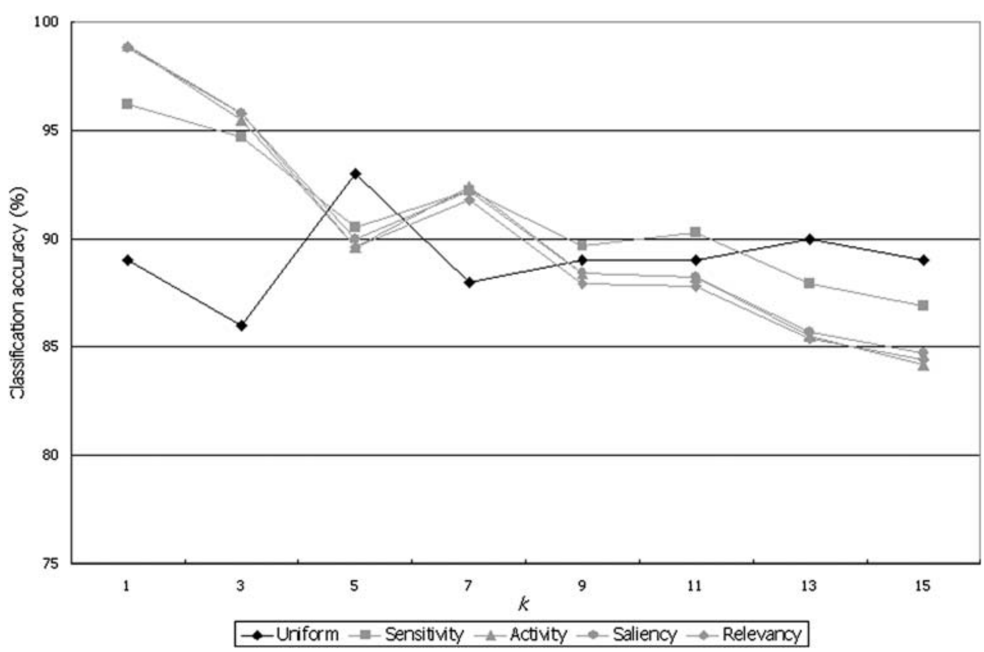


Fig. 5 Personalized advice based on the Chinese five basic elements and skin test

Fig. 6 Classification accuracy of the feature weighting methods—without the KUM



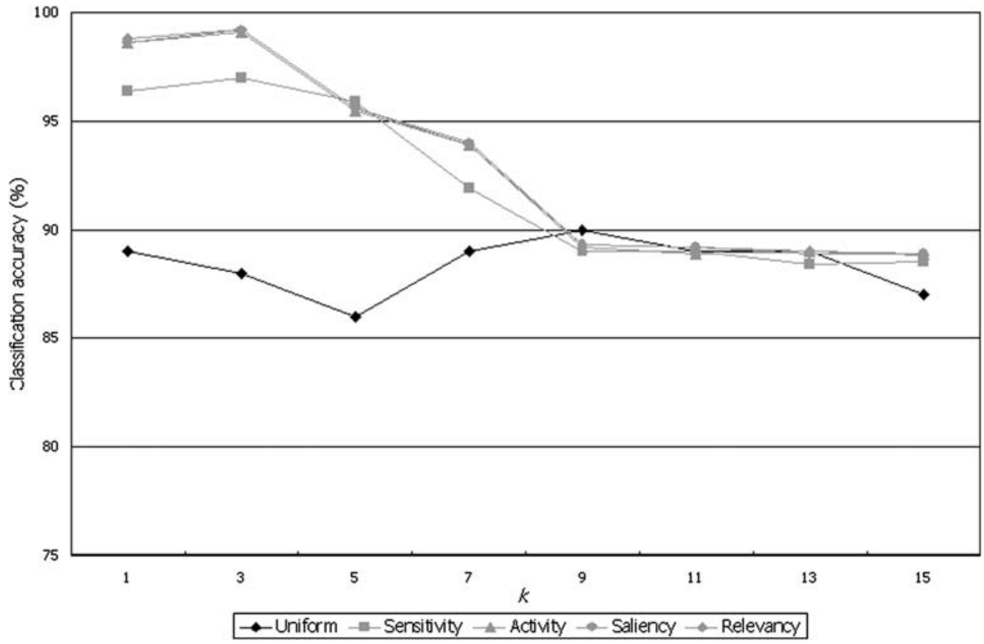


Fig. 7 Classification accuracy of the feature weighting methods—with the KUM

Table 3 Statistics of the classification errors of the feature-weighting methods

k	Uniform	Sensitivity		Activity		Saliency		Relevance	
	Mean	Mean	SD	Mean	SD	Mean	SD	Mean	SD
1	0.110	0.036	0.0177	0.014	0.0171	0.014	0.0171	0.012	0.0168
3	0.120	0.030	0.0163	0.009	0.0110	0.008	0.0103	0.008	0.0113
5	0.140	0.041	0.0087	0.045	0.0084	0.044	0.0096	0.044	0.0084
7	0.110	0.081	0.0228	0.061	0.0185	0.060	0.0169	0.061	0.0191
9	0.100	0.110	0.0000	0.108	0.0063	0.107	0.0094	0.107	0.0094
11	0.110	0.110	0.0000	0.111	0.0031	0.108	0.0063	0.108	0.0063
13	0.110	0.116	0.0069	0.110	0.0047	0.111	0.0056	0.110	0.0047
15	0.130	0.115	0.0052	0.111	0.0031	0.111	0.0031	0.112	0.0042

Figure 7 shows that classification improves in accuracy due to online learning performed by the KUM. The KUM provides new cases to the case base, and the KUM trains the neural network regarding the new cases. The weighting methods, with online learning, exceed those without online learning. This is because supplying new cases continuously prevents the decay of the classification accuracy of the feature-weighting algorithms.

In this experiment, the weighting methods excel the uniform method in every dataset. In particular, when k is less than nine, there is a wide difference in the classification accuracy between the uniform and weighting methods. There are, however, small differences in the classification accuracy between the four feature-weighting methods. In particular, when k is larger than nine, there is little difference between the weighting methods. Little fluctuation in classification accuracy, as k increases, is attributed to the online learning feature of the KUM.

In general, it is difficult to decide which weighting method is the best. In addition, this article doesn't aim to show which option is preferable in which case. However, we suggest that one should test the four methods at the initial development phase and then implement the one with the lowest classification error in the production phase.

Table 3 summarizes the average classification errors and their standard deviations (SD), when implementing both the feature-weighted CBR and the KUM.

The Uniform column shows CBR errors without feature weighting, that is, pure k-NN algorithm. Since we performed an experiment with the uniform method only once, the Uniform column does not contain a standard deviation. The next four columns show mean errors and their standard deviations of the feature-weighting CBR, computed from experiments conducted ten times for each k.

5 Conclusions and future research

We developed a personalized counseling system and applied it to a cosmetic company. We utilized a neural network for weighting symbolic features and a VDM for measuring distances between the values of symbolic features.

The experimental results showed that CANSY classified the five basic elements (metal, wood, earth, water, and fire) with a relatively high level of accuracy. We made use of a decision tree to extract knowledge from experienced beauty counselors and skin measurements. These rules were helpful in identifying skin type and sensitivity of the customers. The personalized counseling system was capable of providing customers with appropriate personalized advice about lifestyle and of recommending appropriate makeup products. This personalization system could improve customer satisfaction.

Although a case base is considered to be a good knowledge representation tool and CBR is a fairly good example of a personalization tool, many problems can occur when using numeric and symbolic attributes together. In our system, answers to questionnaires had symbolic values and the skin test values contained numeric values. Therefore, we utilized the Euclidean distance as a representative distance measurement for numeric features, and the VDM for symbolic features. We have to conduct research about many variants of the VDM which in turn can improve heterogeneous distance functions. Because there are a variety of personalization domains, our framework and methodology have to be verifiable and generalized by other forms of personalization. In addition, in order to overcome the highly subjective nature of the Chinese cosmogony system and consequently of the representations based on it, the current evaluation of the application must be complemented by an evaluation of the customer satisfaction with the system.

References

1. Roberts ML (2003) Internet marketing: integrating online and offline strategies. New York, McGraw-Hill
2. Liao S-H (2005) Expert system methodologies and applications—a decade review from 1995 to 2004. Expert Syst Appl 28:93–103
3. Wettschereck D, Aha DW, Mohri T (1997) A review and empirical evaluation of feature weighting methods for a class of lazy learning algorithms. Artif Intell Rev 11:273–314
4. Ahn H, Kim K, Han I (2006) Global optimization of feature weights and the number of neighbors that combine in a case-based reasoning system. Expert Syst 23:290–301
5. Shin CK, Yun UT, Kim HK, Park SC (2000) A hybrid approach of neural network and memory-based learning to data mining. IEEE Trans Neural Netw 11:637–646
6. Cost S, Salzberg S (1993) A weighted nearest neighbor algorithm for learning with symbolic features. Mach Learn 10:57–78
7. Im KH, Kim TH, Park SC (2005) Hybrid system of case-based reasoning and neural network for symbolic features. Lect Notes Comput Sci 3589:265–274
8. Ball D, Coelho PS, Vuares MJ (2006) Service personalization and loyalty. J Serv Mark 20:391–403
9. Hwang JH, Gu MS, Ryu KH (2005) Context-based recommendation service in ubiquitous commerce. In: Proceedings of the international conference on computational science and its applications
10. Kwon O, Choi S, Park G (2005) NAMA: a context-aware multiagent based web service approach to proactive need identification for personalized reminder systems. Expert Syst Appl 29:17–32
11. Kwon O (2006) The potential roles of context-aware computing technology in optimization-based intelligent decision-making. Expert Syst Appl 31:629–642
12. Bellavista P, Corradi A, Montanan R, Tonineili A (2006) Context-aware semantic discovery for next generation mobile systems. IEEE Commun Mag 44:62–71
13. Leavitt N (2006) Recommendation technology: will it boost e-commerce. IEEE Comput 39:13–16
14. Special issue on personalization (2000) Commun ACM 43:127–158
15. Adomavicius G, Tuzhilin A (2005) Personalization technologies: a process-oriented perspective. Commun ACM 48:83–90
16. Eirinaki M, Vazirgiannis M (2003) Web mining for Web personalization. ACM Trans Internet Technol 3:1–27
17. Ralph P, Parsons J (2006) A framework for automatic online personalization. In: Proceedings of the 39th Hawaii international conference on system sciences
18. Weng S-S, Liu M-J (2004) Feature-based recommendations for one-to-one marketing. Expert Syst Appl 26:493–508
19. Lee H-M, Huang C-C, Kao T-T (2005) Personalized course navigation based on grey relational analysis. Appl Intell 22:83–92
20. Fan H, Poole MS (2006) What is personalization? Perspectives on the design and implementation of personalization in information systems. J Organ Comput Electron Commer 16:179–202
21. Zhang Y, Jiao J (2007) An associative classification-based recommendation system for personalization in B2C e-commerce applications. Expert Syst Appl 33:357–367
22. Li Y, Lu L, Xuefeng L (2005) A hybrid collaborative filtering method for multiple-interests and multiple-content recommendation in e-commerce. Expert Syst Appl 28:67–77
23. Liu H, Motoda H (1998) Feature selection for knowledge discovery and data mining. Norwell, Kluwer
24. Zhang L, Coenen F, Leng P (2002) Formalising optimal feature weight setting in case based diagnosis as linear programming problems. Knowl-Based Syst 15:391–398