Neural Comput & Applic
DOI 10.1007/s00521-016-2707-8

ORIGINAL ARTICLE

Rule-based back propagation neural networks for various precision rough set presented KANSEI knowledge prediction: a case study on shoe product form features extraction

Zairan Li[1,2] · Kai Shi[2] · Nilanjan Dey[3] · Amira S. Ashour[4] · Dan Wang[1] · Valentina E. Balas[5] · Pamela McCauley[6] · Fuqian Shi[7]

Received: 11 October 2016 / Accepted: 3 November 2016

Abstract Nonlinear operators for KANSEI evaluation dataset were significantly developed such as uncertainty reason techniques including rough set, fuzzy set and neural networks. In order to extract more accurate KANSEI knowledge, rule-based presentation was concluded a promising way in KANSEI engineering research. In the present work, variable precision rough set was applied in rule-based system to reduce the complexity of the knowledge database from normal item dataset to high frequent rule set. In addition, evidence theory's reliability indices, namely the support and confidence for rule-based knowledge presentation, were proposed by using back propagation neural network with Bayesian regularization algorithm. The proposed method was applied in shoes KANSEI evaluation system; for a certain KANSEI adjective, the key form features of products were predicted. Some similar algorithms such as Levenberg–Marquardt and scaled conjugate gradient were also discussed and compared to establish the effectiveness of the proposed approach. The experimental results established the effectiveness and feasibility of the proposed algorithms customized for shoe industry, where the proposed back propagation neural network/ Bayesian regularization approach achieved superior performance compared to the other algorithms in terms of the performance, gradient, Mu, Effective number of parameter, and the sum square parameter in KANSEI support and confidence time series prediction.

Keywords KANSEI engineering · Variable precision rough set · Fuzzy set · Back propagation neural networks · Bayesian regularization · Shoe product form design

✉ Fuqian Shi
sfq@wmu.edu.cn

1 Tianjin Key Laboratory of Process Measurement and Control, School of Electrical Engineering and Automation, Tianjin University, Tianjin 300072, People's Republic of China

2 Wenzhou Vocational and Technical College, Wenzhou 325035, People's Republic of China

3 Department of IT, Techno India College of Technology, Kolkata, West Bengal 740000, India

4 Department of Electronics and Electrical Communications Engineering, Faculty of Engineering, Tanta University, Tanta 31111, Egypt

5 Department of Automation and Applied Informatics, Aurel Vlaicu University of Arad, 310130 Arad, Romania

6 Department of Industrial Engineering and Management Systems, University of Central Florida, Orlando, FL 32825, USA

7 College of Information and Engineering, Wenzhou Medical University, 325035 Wenzhou, People's Republic of China

1 Introduction

Artificial intelligence is a new technology to extend theory, method, technology and application system for simulating, extending and expanding human intelligence. Artificial intelligence is also a branch of computer science, which attempts to understand the essence of intelligence, produce a new kind of human intelligence and can be a similar way to respond to intelligent machines. It has more applications in the field of robotics, speech recognition, image recognition, natural language processing and expert systems. The research of uncertainty reasoning has made important contributions to a series of fundamental theoretical problems of artificial intelligence, such as non-monotonic reasoning, heuristic reasoning and planning learning.

User-friendliness, manufacturability and ecological considerations complicate the product design in modern society. Through the whole design stage, the feasibility and the individuals' satisfaction become the most important issue. Since KANSEI engineering refers to the scheme that reflects the customer's response regarding a specific product and its parameters to support future product design, the KANSEI extracts the individuals' response and feeling through the evaluation process using adjectives, words or sentences. It translates the consumer's feeling and image for a product into design parameters [38]. KANSEI engineering is regarded as an effective tool for realizing ergonomics and consumer-oriented technology [10, 39] and widely utilized in electrical appliance and automotive applications in high level [1]. KANSEI mapping between human feeling and product forms is often a black-box.

Many uncertainty inference tools for KANSEI engineering including the accuracy in representing the consumer's KANSEI, the KANSEI translation into product design, and the creation of KANSEI-orientated system naturally are under development. Thus, researches are interested with proposing some mathematical tools for precisely describing the mapping, such as soft computing (SC), which defines an emerging methodology for reasoning and learning human needs, including fuzzy set, rough set, genetic algorithm, neural networks, and statistics inferences [21]. Wang [58] proposed binary linear classifier for two or more possible classes called support vector regression (SVR) combined with gray system theory (GST) for hybrid design expert system. Fu [50] applied quantification theory in automobile head-up image display for drivers' KANSEI. Furthermore, artificial KANSEI collaboration [55], general multi-attribute target-oriented evaluation function [18] and clustering adjectives [20] were introduced subsequently. Recently, researchers apply KANSEI on other industries more than product design only, such as post-disaster evaluation [22], service industries [9] and medical imaging experiments [17]. Particularly, Petiot founded an auditory parameter method for emotion measurement which linked the customer's emotional responses, the product's geometrical features and the product's semantic [33].

Since rough set is "a formal approximation of a crisp set in terms of a pair of sets" which provides the lower and the upper approximation of the original set, thus, the KANSEI knowledge from imprecise design information was effectively extracted using the rough set. Zhai et al. [63, 64] introduced a systematic approach in affective design. Consequently, consumer affective responses can be captured exterior features so-called preference using rough set [37]. In addition, the rough set combined with the particle swarm optimization (PSO) [53] based adaptive neuron fuzzy inference system (ANFIS) [23] was proposed. Probabilistic rough set based on Bayesian decision procedure [29] and attribute reductions to increase computation complex [7] were introduced subsequently. In particular, variable rough set for KANSEI was proved in KANSEI engineering to exact reflection of the knowledge equivalence relations [62]. Huang and Li [19] introduced the fuzzy set theory (FST) and variable precision rough set for multiple-criteria decision-making based on statistical approach.

Lin et al. [32] focused on hierarchical level-search procedure using rough set and decision rules. The back propagation neural network (BPNN) proved its efficiency for easier knowledge presentation pre-processing. In 2008, Ushada and Murase [56] extracted features from human preferences using a Bayesian belief network (BBN), where reasoning based on conditional probability table (CPT), which is a successfully application. Hidden Naïve Bayes [13, 26, 61], Bayesian rough set model (BRSM) [65], and high-dimensional text data using Bayesian network [24] were also developed, whereas value weighting method classification learning algorithms was regarded as a feasible method in KANSEI multi-decision-making issue [27, 66].

In KANSEI research, rule-based presentation was concluded to acquire exact and accurate knowledge through the formal rule presentation with confidence and support value [40, 44, 48], even for large datasets [41]. Kernel principal component analysis [45], fuzzy association rules mining [43], rule-based genetic algorithm [4] and multi-objective genetic algorithm [15] were widely utilized for affective effect of design patterns. Associate rule-based learning technology deals with nonlinear issues in a superior way, where the associate rule mining can easily determine the interesting relational association rules [11], and specially by using fuzzy set theory [54]. Some studies established that the product form, the customer's impression, design parameters or the perceptual values for the design can be linked by using the artificial neural network (ANN) [31]. Furthermore, some artificial intelligent tools were also involved, such as the principle component analysis (PCA)-based neural network, partial least squares [49] for nonlinear dataset, genetic algorithm combined with fuzzy neural network for road traffic control actions [12], daily worker evaluation model for KANSEI parameters using ANN [57], the PSO algorithm combined with the BPNN for training [52], Bayesian network-based BPNN and automatic credit scoring systems using back propagation algorithm [8].

As focused research in this paper is acutely relative to data mining in knowledge discovery, in general, data mining is a step in database knowledge discovery which refers to the process of hidden information in a large amount of data through algorithms. Data mining usually achieves these goals related to computer science through statistical, online analysis and processing, information retrieval, machine learning, expert system (relying on past experience rules), pattern recognition and other methods.

Classification is used to select several categories from training set, and the data mining classification technique is applied to the training set. The classification model is established, and the classified data is classified [46]; estimation is similar to classification, except that the classification describes the output of discrete variables and the output of continuous values of the estimation process; the classification of data mining is classified by a certain number, and the amount of estimation is uncertain. In general, the prediction is affected by classification or valuation, that is, by classifying or estimating the model, which is used for the prediction of unknown variables. In this sense, the prophecy is not necessarily divided into a single class. Predicting purpose is to predict future unknown variables [3]. This prediction is time to verify, that is, after a certain period of time before we know how much accuracy of the prediction. Clustering is a grouping of records and a similar record in a cluster. The difference between clustering and classification is that aggregation does not depend on pre-defined classes and does not require training sets. There are some significant advantages we need to concern that are accuracy, speed, robustness, scalability and interpretability in data mining for knowledge discovery [14, 30].

Consequently, the artificial intelligence technologies can be used to create an inference and prediction system; however, the linear classifier (LDC) for the linear combination of features value, K-mean neural network (K-NN) for clustering by the nearest neighbors, the Gaussian mixture model (GMM) as normal distribution model, and the decision tree algorithms suffer from local limitation and cannot be suitable for KANSEI engineering issues. In addition, a statistical Markov model named hidden Markov model (HMM) was proved for speech affect only in an effective way. In the present work, the proposed algorithm was applied for shoe design research to show the effectiveness with comparative analysis. For this special industry, in 1996, Barton first introduced neural network technology for shoe insole materials [2], and semantic differential [34] and topological optimization [42] were also introduced for shoe design research. Recently, high-heeled shoes, heel base size (HBS), the trajectory of the center of pressure (COP), maximal peak pressure, pressure time integral, contact area, and perceived stability were analyzed to the joint torque and muscle activity and induce lower limb problems [35].

In the current work, the BPNN and Bayesian regularization for KANSEI inference were applied. Artificial neural network is a system with learning ability and can develop knowledge, more than the level of knowledge of the original designer. Usually, its training can be divided into two kinds, one is supervised or supervised learning, it uses the given sample standard classification or imitation; the other is a non-supervised learning or unsupervised learning, it only requires learning or some rules, specific learning content with the system in the environment (i.e., the input signal varies depending on the situation) [6]; the system can automatically find environmental characteristics and laws, which is more similar to the function of the human brain. BP (propagation back) neural network, which is the learning process of back propagation error back propagation algorithm, is composed of two processes: the forward propagation of information and the back propagation of error. The input layer neurons are responsible for receiving the input information from the outside world and transfer to the middle layer neurons; the middle layer is the layer responsible for internal information processing and information transform, according to the information demand, the middle layer can be designed as a single hidden layer or multi-hidden layer structure; the last hidden layer to the output layer neurons that transfer the information is further processed after the completion of a forward propagation process of learning, from the output layer to the external output information processing [5]. BP neural network also is a kind of error back propagation training algorithm for the multilayer feed forward network, and the neural network is one of the most widely used models as it can learn and store a large number of input–output model mapping, without revealing the mathematical equation describing the relationship of the mapping. Its learning rule is to use the gradient descent method to adjust the weight and threshold of the network, which makes the error square and minimum of the network. The topology structure of BP neural network model includes input layer (input), hidden layer (layer hidden) and output layer (layer output). The artificial neural network has the ability of self-adaptation and self-organization. In the process of learning or training to change the synaptic weight value, in order to adapt to the requirements of the surrounding environment, the same network can have different functions because of different learning methods and contents [28, 36].

Neural networks can be used for classification, clustering, prediction and so on. Neural networks need to have a certain amount of historical data, and through the training of historical data, the network can learn the hidden knowledge in the data. In your question, first of all to find some of the characteristics of some of the problems, as well as the corresponding evaluation data, using these data to train the neural network.

The Bayes inference and Bayes regularization were adapted in the current work. The BP neural network was created for acquiring the KANSEI knowledge using Bayes regularization and variable precision rough set. For multi-decision problems, Bayes model was proved feasible, while KANSEI knowledge acquisition also is a multi-decision problem. A rule-based presentation way for neural network inference was developed as previously the neural network was used to the treat natural language [47]. Time series forecasting was proposed in the existing work, where using classic BPNN will be falling into the local minimum point, but there will be some measurement for avoiding, such as adaptive neural network [51, 59]. Moreover, for precisely describing the KANSEI knowledge, improved rouge set presentation based on the previous work was proposed in [16].

In rough set theory (RST), there is a close relation between rough approximation and fuzziness of the set, uncertainty and rough membership. In order to derive non-strong decision rules for decision probability estimation, we should consider the degree of overlap. Ziarko extended the idea of rough membership function and proposed a probabilistic method of rough set-variable precision rough set [60]. Variable precision rough set is an important expansion of rough set which deal with partial classification based a probability value, and it also reacts the portion of correct classification in the target dataset. One of the significant contributions of variable precision rough set is to make attribution reduction through a priori algorithms, and the proposed method was also discussed in our previous study [16]. The BP neural network for knowledge discovery applied in this paper is a new contribution in KANSEI research filed. The traditional knowledge base has many disadvantages such as difficulty in knowledge acquiring, complex expression and easy to produce combinatorial explosion for reasoning. Based on the knowledge-based construction method of neural network, especially combining with the structure and training algorithm of BP network, the knowledge acquisition is proved as simple in expression and clearly in reasoning logic [25, 28].

The organization of the remaining sections is as follows. Section 2 presents the modeling for variable precision rough set KANSEI presentation, Bayesian inference and regularization methods, rule-based neural network prediction, and the case study on shoe design applying the proposed methods to show how the modeling runs in KANSEI issues of shoe design and the prediction model for quick response the market change. Section 3 addresses the results and comparative analysis. Section 4 involves the conclusions of the present study.

2 Methodology

2.1 The KANSEI knowledge presentation and inference

From recognition physiological aspects, the KANSEI recognition relies on personal observation of customers. In addition, natural language can be normally used for describing the feeling on the form features of product. However, simple words or adjectives are much more effective, such as "simple", "fashion". This is called semantic image which consisted in the KANSEI space. The mapping between the words and the images is constructed in the current work as illustrated in Fig. 1, where the questionnaire system was founded to extract form features (S_n) of shoe and the image adjectives (K_n).

In Fig. 1, the KANSEI knowledge was formalized by direct graph, the start node represents the product form's features, while the end node is the KANSEI image, and the arc is the rule set that supports to make the conclusion (evaluation process). In addition, some support is required to evaluate the inference's reliability as "IF E THEN" (LS, LN) H P(H), in which LS and LN are the sufficiency and necessity of evidence E and P(H) is prior probability of the conclusion. Since the KANSEI knowledge rule discovery is influenced by the experts, so in order to solve this problem, LS and LN are configured by field expert. The proposed approach analyzed the significance of knowledge rule and the necessity of knowledge by the essence of KANSEI knowledge to define LS and LN.

The perceptual cognition of the household is not only reflected in the identification of the local modeling features, but also the overall product characteristics and the structure of the product. Hence, as a kind of KANSEI knowledge inference, this is also a kind of multi-evidence synthetic calculation, which can be considered knowledge discovery process of the different form features of the product. In the current study, the customer's preference for the whole model was acquired by investigating the KANSEI reflection of the local shape of the product. Hence, the probability function is introduced as follows:

$$O(x) = \frac{P(x)}{1 - P(x)}, \quad P(x) = \frac{O(x)}{1 + O(x)}, \tag{1}$$

where $P(x)$ and $O(x)$ is uniform monotone, where

$$P(H|E) = \frac{LS \times P(H)}{(LS - 1) \times P(H) + 1} \tag{2}$$

If there are n evidences to support the same conclusion and the premise of each evidence will appear, thus, it is only essential to calculate $O(H|E_i)$ of each evidence. Afterward, the following evidence synthetic probability formula is used:

$$\begin{aligned} O(H|E_1, E_2, \cdots, E_n) = {} & \frac{O(H|E_1)}{O(H)} \times \frac{O(H|E_2)}{O(H)} \\ & \times \cdots \times \frac{O(H|E_n)}{O(H)} \times O(H) \end{aligned} \tag{3}$$

$$P(H|E_1, E_2, \cdots E_n) = \frac{O(H|E_1, E_2, \cdots E_n)}{1 + O(H|E_1, E_2, \cdots E_n)} \tag{4}$$

In data mining, one of the important and powerful classification techniques is Bayes network that has the smallest

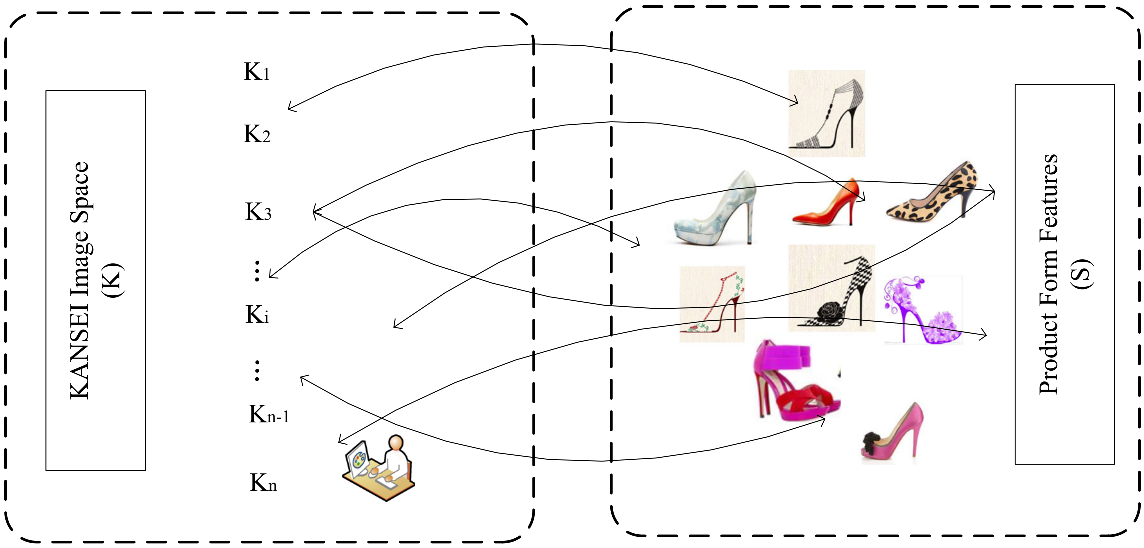

Fig. 1 KANSEI mapping between semantic adjectives and product form features

error rate. In addition, the rough set theory is a new theory which is applied for dealing with inaccurate, uncertain and incomplete data and information. It provides a significant tool for the analysis, reasoning and expression of uncertain information and data. In order to enhance the adaptability of rough set, the variable precision rough set model by Ziarko is proposed. In addition, the Bayesian rough set model based on probability representation is employed. In this paper, a new Bayesian inference method was constructed to calculate the various parameters of the generative KANSEI knowledge rules by using the definition of sample distance measure and the concept of variable precision rough set. This method not only overcomes the shortcomings of the past technologies, but also makes the reasoning and access to the knowledge of the human closer to human's general recognition.

2.2 Definitions

Now we give definitions on feature coding and variable precision rough set for attribution reduction, in which we focused on knowledge presentation using Definitions (1–6). Definitions (7–10) are for variable precision rough set [38, 39]; Definitions (11–15) are for fuzzy associate rule mining which we used in our previous study [16].

Definition 1 Let the product is consisted of different form items S_i (supposed that each item is not interference), each item involved a number of different form features S_{ij}, and binary coding method was applied for these features.

Definition 2 The distance between two samples, i.e. the difference of preferences between customer E_i and E_j on form features, is defined as:

$$g(i,j) = \sum_{k=1}^{m} \| S_{ki} \quad XOR \quad S_{kj} \|, \tag{5}$$

where $\| S_{ki} \quad XOR \quad S_{kj} \| = \begin{cases} 0 & S_{ki} \quad XOR \quad S_{kj} = 0 \\ 1 & S_{ki} \quad XOR \quad S_{kj} \neq 0 \end{cases}$, S_{ki} and S_{kj} presents the selection on class k by customer E_i and E_j; respectively.

Definition 3 The probability of S_{ij} is calculated by:

$$P(S_{ij}) = \frac{C(S_{ij})}{n}, \tag{6}$$

where $C(S_{ij})$ presents the frequent of form feature S_{ij} in column i.

Definition 4 The highest frequent form feature is selected by customer, named as typical preference feature, which is calculated by:

$$S_i^{\mathrm{Max}} = \{S_{ip} | \forall j, P(S_{ip}) \geq P(S_{ij})\} \tag{7}$$

Definition 5 For the inconsistency between the customer's preference and the features of the typical preference, the deviation degree of the customer's KANSEI evaluation $A = \{S_i^{\mathrm{Max}}\}$, $i = 1, 2, \ldots m$ is defined as:

$$g(i, \max) = \frac{1}{m} \sum_{k=1}^{m} \| S_{ki} \quad XOR \quad S_k^{\mathrm{Max}} \|, \tag{8}$$

where $0 \leq g(i, \max) \leq 1$.

Definition 6 Likert scale is applied for KANSEI evaluation, and the average Likert value is calculated by:

$$AL = \frac{1}{n \cdot Len\_L} \sum_{l=1}^{n} Lik(E_l), \tag{9}$$

where Len_L is the length of Likert scale. Along with these definitions, other ones are reported as follows for the rough set and variable rough set definitions.

Definition 7 Rough set: let $I = (U, A)$ be an information system, where U is a non-empty set of finite objects and A is a non-empty, finite set of attributes as $a : U \rightarrow V_a$ for every $a \in A, V_a$ is the set of values that attribute a may take. The information table assigns a value $a(x)$ from V_a to each attribute a and object x in the universe U. In addition, with any $P \in A$ there is an associated equivalence relation, which is given by:

$$IND(P) = \{(x, y) \in U^2 | \forall a \in P, a(x) = a(y)\} \tag{10}$$

The relation $IND(P)$ is called a P-indiscernible relation. The partition of U is a family of all equivalence classes of $IND(P)$ and is denoted by $\frac{U}{IND(P)}$. If $(x, y) \in IND(P)$, then x and y are indiscernible (or indistinguishable) by attributes from P.

Definition 8 The upper approximation set of variable precision rough set is defined as:

$$POS_\alpha(E) = \cup\left\{E_p | g(p, \max) > \frac{AL - \alpha}{1 - \alpha}, 0 \le \alpha \le 1 - AL\right\} \tag{11}$$

Definition 9 The lower approximation set of variable precision rough set is defined as:

$$NEG_\alpha(E) = \cup\left\{E_p | g(p, \max) < \frac{AL}{1 - \alpha}\right\} \tag{12}$$

Definition 10 The boundary of variable precision rough set is defined as:

$$BND_\alpha(E) = POS_\alpha(E) \backslash NEG_\alpha(E) \cup \{E_p | g(p, \max) \tag{13}$$

Obviously,

$$BND_\alpha(E) \in \left(\frac{AL - \alpha}{1 - \alpha}, \frac{AL}{1 - \alpha}\right), \quad 0 \le \alpha \le 1 - AL. \tag{14}$$

Definition 11 KANSEI knowledge is presented by rule-based system which is formalized as:

$$IF \quad S_{ij} \quad THEN \quad (LS, LN) \quad A \quad CER(S_{ij}A) \tag{15}$$

Definition 12 The knowledge is much more precision, while LS is bigger, where $LS = 1$ shows that knowledge is not real. Here, LS is defined as follows:

$$LS = 1 + ||BND_\alpha(E)||^{-1}, \tag{16}$$

where $||BND_\alpha(E)||$ is norm of the boundary, it can be the interval in Eq. (14). Furthermore, LN is the support for negative of the evidence, while $LN = 1$ denotes real evidence, $LN > 1$ shows not real evidence, so it is regarded that $LN \le 1$ and LN close to 0 show the conclusion is not real based on the evidence absence. So LN can be defined as:

$$LN = \frac{||D|| - ||BND_\alpha(E)||}{||D||}, \tag{17}$$

where $||D||$ is the total number of the evidence in universe.

Definition 13 Reliability of KANSEI evaluation of the customer E_l is defined as:

$$CER(E_l) = 1 - \frac{|A_l - AL|}{Len\_L} \tag{18}$$

Definition 14 $CER(E_l)$ can be defined as the reliability of Product form A, which is given by:

$$CER(S_{ij}, A) = \sum_l \{CER(E_l) | S_{ij} \in E_l\}. \tag{19}$$

Definition 15 The typical form feature's initial probability is defined as:

$$P(A) = \frac{1}{m} \sum_{i=1}^{m} CER\left(A, S_i^{\mathrm{Max}}\right). \tag{20}$$

2.3 Bayesian inference and regularization

The descriptive formal of the Bayesian inference uses the following definitions:x is a data point (vector) in universe, θ is the parameter vector of the x's distribution as $x \rightarrow p(x|\theta)$,$\alpha$ is the hyper parameter vector of the parameter as $\theta \rightarrow p(\theta|\alpha)$, X is a set of n observed data points as x_1,x_2,..., x_n and $\tilde{x}$ is a new data point whose distribution is to be predicted. The prior distribution of the parameters before any data is observed, which given by $p(\theta|\alpha)$ is not easy to be determined. Thus, in order to apply the Bayesian inference, a posterior distribution for updation using newer observations is used, which can be viewed as a function of the parameters that is written as $L(\theta|X) = p(X|\theta)$. The thoughts of regularization are to reserve all variables under decreasing parameter θ avoiding an eigenvector's over-influence. Based on the inference mechanism of Bayesian statistics, we introduce the regularization process in more details as follows.

In previous studies, in general, we use maximum likelihood estimation on parameter θ to get minimum value of cost function as,

$$\theta_{ML} = \arg\max_\theta \prod_{i=1}^{m} p\left(y^{(i)} | x^{(i)}; \theta\right) \tag{21}$$

And θ is unknown random variables, so before we set the training of training, theta may obey a certain distribution of $p(\theta)$, we call the prior probability (prior distribution). For a new training set $S = \{(x^{(i)}, y^{(i)})\}$, $i = 1, \ldots, m$. We can use Bayesian formula to get poster distribution of θ as,

$$p(\theta|S) = \frac{p(\theta|S)p(\theta)}{P(S)} = \frac{\prod_{i=1}^{m} p(y^{(i)}|x^{(i)},\theta)p(\theta)}{\int\limits_{\theta} \prod_{i=1}^{m} p(y^{(i)}|x^{(i)},\theta)p(\theta)d\theta} \tag{22}$$

where $p(y^{(i)}|x^{(i)},\theta)$ is decided by the learning model and to predict x using poster distribution of θ, continuously, we have,

$$p(y|x,S) = \int\limits_{\theta} p(y|x,\theta)p(\theta|S)\mathrm{d}\theta \tag{23}$$

And for y, we have,

$$E[y|x,S] = \int\limits_{y} yp(y|x,S)\mathrm{d}y \tag{24}$$

In fact, it is difficult to calculate the post-distribution of θ for that θ is multi-dimensional variable which we need to integrate formula (22). So an approximate method was adopted to replace (22) called maximum a posteriori as follows,

$$\theta_{MAP} = \arg\max_{\theta} \prod_{i=1}^{m} p(y^{(i)}|x^{(i)},\theta)p(\theta) \tag{25}$$

And (25) is only multiple prior distribution of θ, in practice, $\theta \sim N(0,\tau^2 I)$ so the Bayesian MAP estimate has better performance on fitting than maximum likelihood estimation.

For some data, Bayesian regularization is essential before training. In fact, the Bayes' rule is a conditional probability and normally using logistic regression for maximal likelihood estimation. On the basis of the formal parameters and calculation methods, the research constructs the Bayesian inference network as shown in Fig. 2.

The network illustrated in Fig. 2 includes the root node which represents the perceptual image semantics that is the conclusion of the knowledge rule. At the same time, according to the relative independence principle, the sub-nodes represent the typical preference characteristics, namely, the evidence of knowledge rule. If the design scheme is composed of the user's perceptual

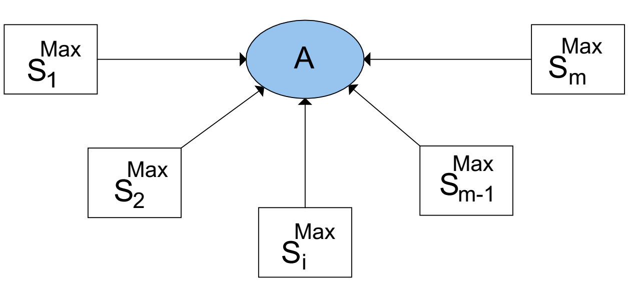

Fig. 2 Bayesian inference network

preference, then the perceptual knowledge rule set can be given by:

$$\begin{array}{llll} R1: IF & S_1^{Max} & THEN & (LS_1, LN_1) \quad A \\ R2: IF & S_2^{Max} & THEN & (LS_2, LN_2) \quad A \\ & \ldots & & \\ Rm: IF & S_m^{Max} & THEN & (LS_m, LN_m) \quad A \end{array}$$

At this point, perceptual knowledge rule acquisition is transformed into a typical multi-evidence synthesis of Bayesian computation problem. The problem solving can be calculated by the theory of the Bayesian multiple evidence synthesis, and the credibility of the perceptual knowledge is obtained.

2.4 Fuzzy associate rule mining and back propagation neural networks

Let KANSEI adjectives set be $CF = \{Z_1, Z_2, \ldots, Z_k\}$ and form features set be $PF = \{\{A_{j1}, \ldots, A_{jk}\}|j = 1, \ldots, n; k = i_1, \ldots, i_p\}$, where $\{A_{1j}|j = 1, \ldots, i_1\}$ is the first item of the product forms, whereas $\{A_{pj}|j = 1, \ldots, i_1\}$ is the pth number of form features n, each form feature is consisted of some form elements. Thus, the product evaluation system based on web is investigated by the customers. The evaluation score is given by the customers on every elements of the form feature. This score indicates that the product form feature can be represented by the specific KANSEI evaluation process. The mapping relationship between KANSEI information and form features can be found by fuzzy mining association rules $[Z_k \Rightarrow A_{ij}]$. The whole mining and training framework of the proposed methods are illustrated in Fig. 3. There are two dataset we need to concern that one is the field knowledge and another is form features; the relationship between these two set need to be constructed in this research; these two set were under a questionnaire system by using scoring system. Hereafter, the transaction database was acquired. The database was reorganized as rule set presented like IF–THEN which was formalized in Sect. 2.3. BNPP was used for prediction and FADM was used for knowledge reduction. FADM algorithm was introduced in Sect. 2.4.1 in more details. The basic blocks in Fig. 3 are discussed as follows.

2.4.1 Fuzzy associate rule mining

In the fuzzy relational data mining algorithm (FADM), the algorithm input includes: (i) the training data of n, where each training data have m attributes; (ii) each attribute corresponding to the attribution function; and (iii) the minimum support. In addition, the algorithm output is a collection of fuzzy association rules. Identify the FADM parameters as follows: n is the record number

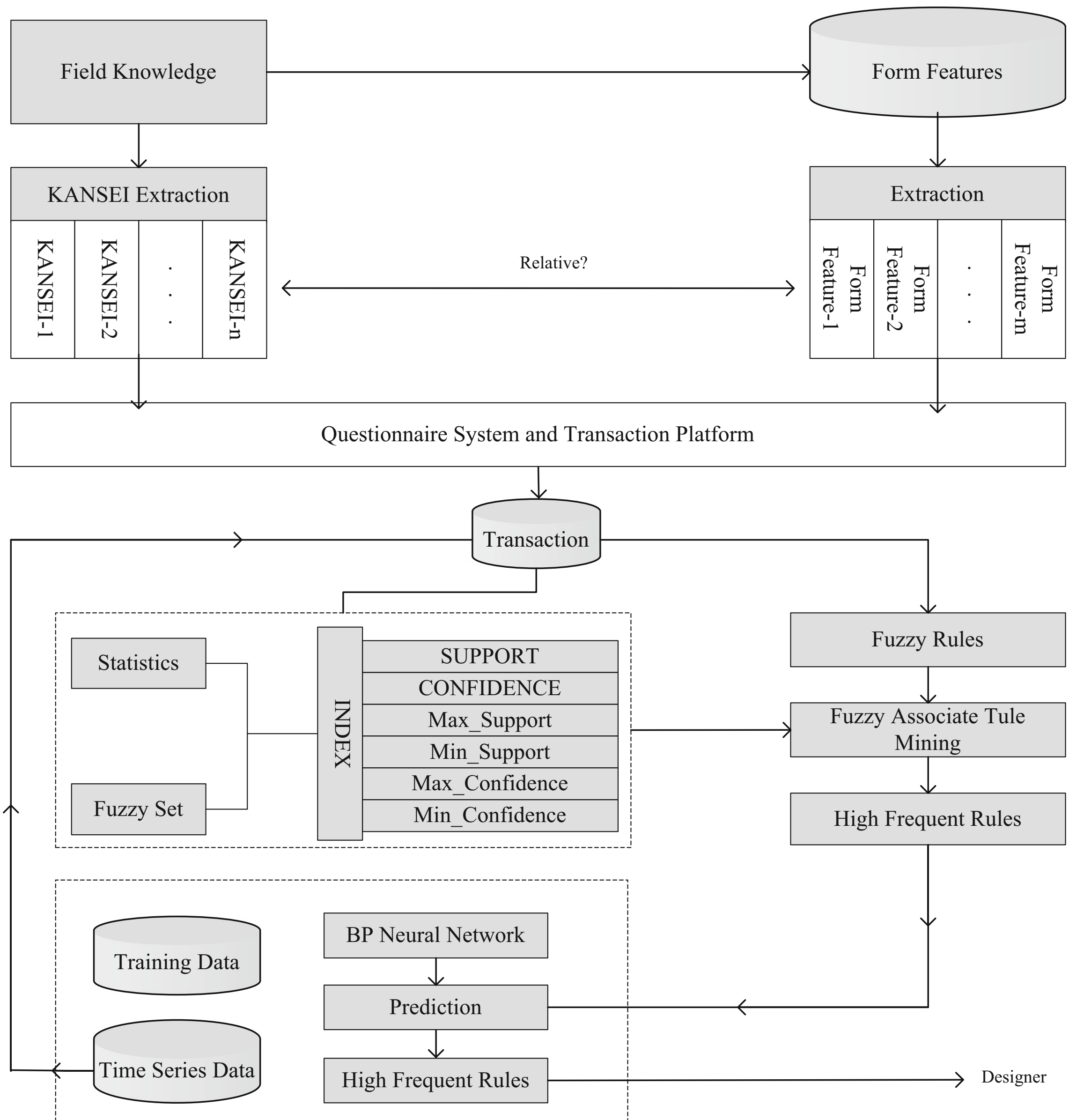

Fig. 3 The framework of FADM and BPNN

of transaction dataset $Count(D)$, m is the number of attributes, A_j is the jth attribute $1 \leq j \leq m$, $|A_j|$ is the KANSEI items for each attribute, R_{jk} is the kth fuzzy interval of the jth attribute $1 \leq k \leq |A_j|$, $D^{(i)}$ is the ith record, $v_j^{(i)}$ is the jth value of $D^{(i)}$, $f_j^{(i)}$ is the fuzzy set from $v_j^{(i)}$, $f_{jk}^{(i)}$ is the membership of $v_j^{(i)}$ on R_{jk}, $count_{jk}$ is the sum of $f_{jk}^{(i)}$, $i = 1, 2, \ldots, n$, α_F is the minimum support, C_r is the r rank candidate set, and L_r is the r rank high-frequency set. The FADM algorithm can be divided into nine steps as follows.

Algorithm: FADM algorithm

Start

(1) ***Turn*** the original data in the database into a fuzzy set of perceptual words because the FADM algorithm is the mining of the association rules of fuzzy sets. Thus, every record must be expressed using the following formula:

$$A_{ij} = \{\frac{f_{j1}^{(i)}}{R_{j1}} + \frac{f_{j2}^{(i)}}{R_{j2}} + \ldots + \frac{f_{jk}^{(i)}}{R_{jk}}\} \tag{26}$$

(2) ***Count*** all KANSEI evaluated form items for each attribute $Count_{jk} = \sum_{i=1}^{n} f_{jk}^{(i)}$.

(3)***Define*** the support as $Support_{jk} = \frac{Count_{jk}}{n}$

(4) ***Initialize*** the candidate set C_1

(5) ***Set*** the threshold α_F for all candidate set on support issue, if true, then the items will be added to 1-rank frequent-item sets (L_1), the formula is $L_1 = \{Support_{jk} \geq \alpha_F, 1 \leq j \leq m\}$, where the threshold will be arranged on different value based different stage.

(6) ***If*** L_1 is not empty, then go to step(7), else end

(7) ***Run*** an improved Apriori liked algorithm, set $r = 1$, joined all high frequent-item sets L_r to generate candidate set C_{r+1} , it may not be joined if both items of L_r have the same attributes while joining the items. For joining operator on high frequent item set L_r , the following steps are followed to generate C_{r+1} :

a) ***subsection*** for L_r joining as

$$\mu_{A \cap B}(x) = \min\{\mu_A(x), \mu_B(x)\}, \forall x \in X \tag{27}$$

b) ***let*** s be the $(r+1)^{th}$ item set of candidate set C_{r+1} , which includes items $S_1, S_2, \ldots, S_{r+1}$ and fuzziness on these items was calculated by:

$$f_s^{(i)} = f_{s_1}^{(i)} \wedge f_{s_2}^{(i)} \wedge \ldots \wedge f_{s_{r+1}}^{(i)} = \min\{f_{s_j}^{(i)} \mid j = 1, \ldots, r+1\} \tag{28}$$

c) ***set*** $Count_s = \sum_{i=1}^{n} f_s^{(i)}$

d) ***set*** $Support_s = \frac{Count_s}{n}$

(8)***If*** $Count_s > \alpha_F$, then generate L_{r+1} .

(9) ***If*** L_{r+1} is empty, then go to the end, else set $r = r + 1$, go to steps (7) and (8)

Stop

2.4.2 Back propagation neural network training with Bayesian regularization

In the current work, the high-frequency rule set generated by time period data is used as the training data of the BPNN. Through the support and confidence of each rule inputting, some indexes are trained, and the rules are adjusted. The whole operation process can be divided into the following steps:

Step 1 Determine the size of the time segment, such as seasonal, half year, or one year of the products' transaction data, denoted as $Data(i)$, $i = 1, 2, 3, \ldots$

Step 2 Mining fuzzy association rules $Rule(i)$ from $Data(i)$, and calculate relevant attributes of each rule's support (S) and confidence (C) store to the knowledge database.

Step 3 When the database is accumulated to $Data\,(i + 1)$, network (BPN) is calculated in every rule in

Time $(i+1)$, and the support and confidence rules are input variables. Then, the obtained support and confidence rules are considered the output variable. Afterward, using the trained BPNN to predict the future support and confidence of each rule in the knowledge base is finished. If the support and confidence of the prediction are greater than or equal to the given threshold values, the quality of the rules is expected to be retained; otherwise, if the support and confidence of the prediction are less than the given threshold value, then rule will be deleted.

Step 4 The $Rule(i+1)$ and $Rule(i)$ are integrated into the knowledge database.

Step 5 For $i = i + 1$ and go to Step 3

Figure 4 demonstrates the BPNN training system based on nonlinear autoregressive.

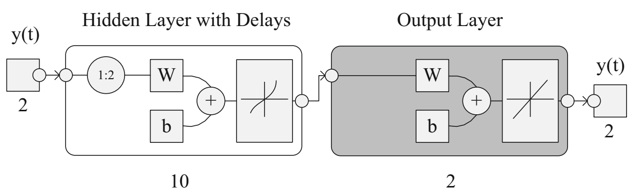

Fig. 4 Back propagation neural networks training system based on nonlinear autoregressive

2.5 Case study

2.5.1 Description of product database

Let form features of high heel shoes' set be $PF = \{\{A_{1,1},\ldots,A_{1,15}\},\ldots,, \{A_{9,1},\ldots A_{9,15}\}\}$ and 9 items set are length of the height, tangent vector for back binding, high heel length, angle for quarter, angle of up sole, support board length, front angle-1, vamp length, inclination length-1}. So, the total number of extracted features is 135. Let KANSEI adjectives set be $KA = \{Z_1, Z_2, \ldots, Z_6\}$. Using WEB (web-based questionnaire) system for product KANSEI evaluation is to give the product of the KANSEI adjectives as "fashion", "lovely", "mature", "scientific", "technological" and "exquisite". The evaluation of the $[1, 5]$ interval is used to indicate the degree of the KANSEI adjectives of the feature modeling element. Through the analysis of women's high-heel shoe, the over whole products and the extracted forms' feature set PF are shown in the following Figs. 5 and 6.

In addition, Table 1 reported the transaction database in the year 2014 for the user's evaluation score, where a total of 9 typical modeling features are described, and a description of the specific value is given.

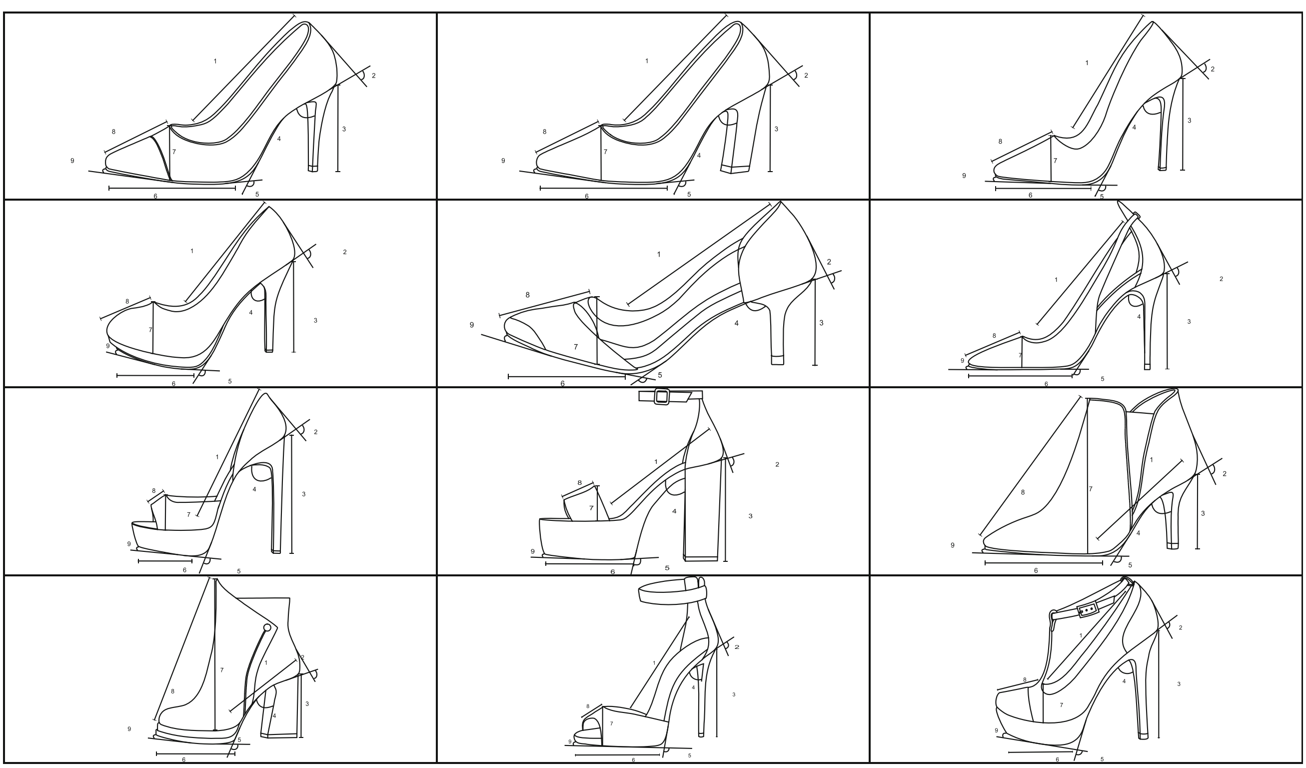

Fig. 5 The typical form features of high heel shoe

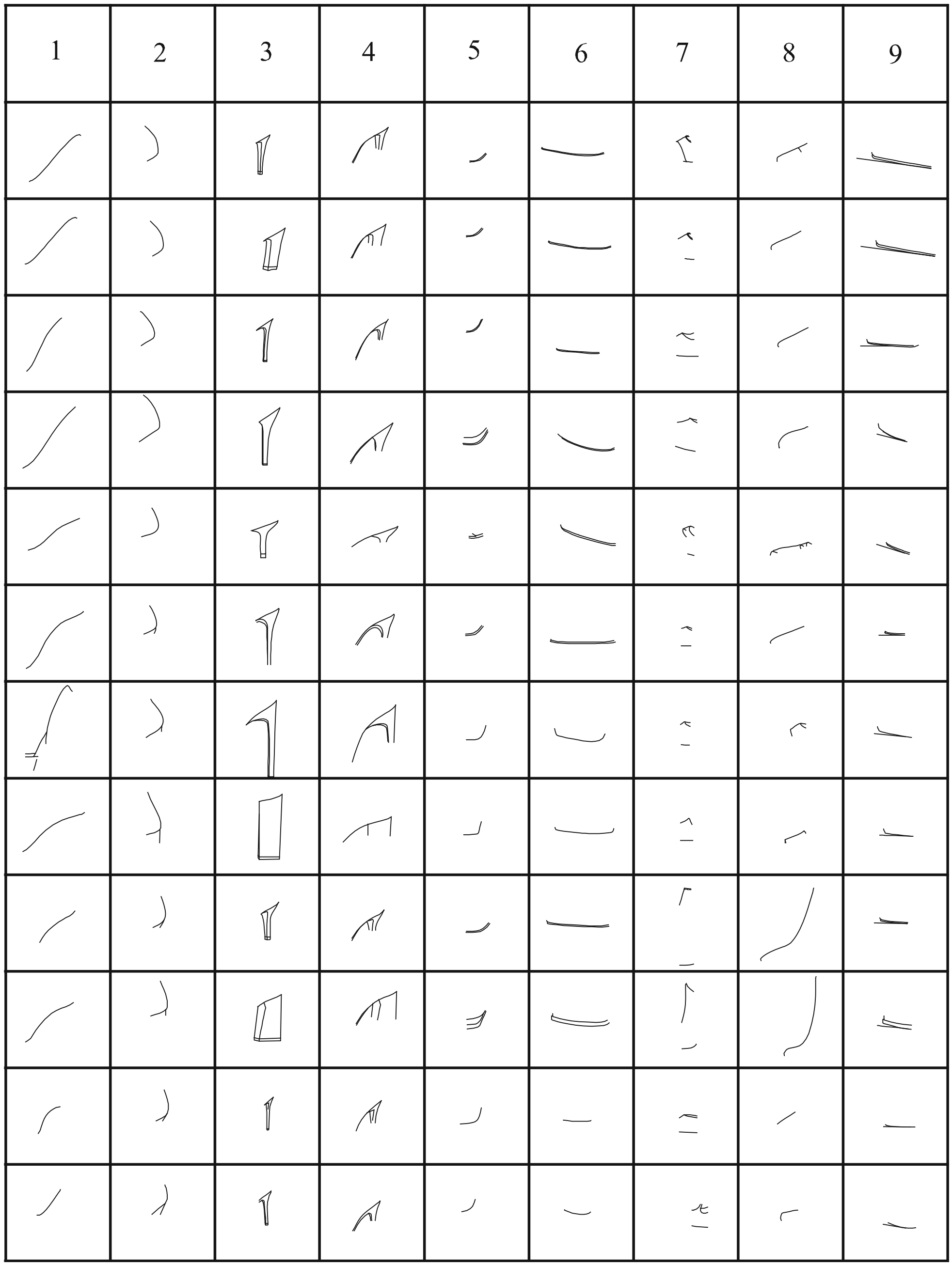

Fig. 6 The forms' features of high heel shoes-PF set

Table 1 Transaction database in 2014

TID	A_{11}	A_{12}	A_{13}	A_{14}	A_{15}	A_{16}	A_{17}	A_{18}	A_{19}
001	Z_1, Z_3, Z_5	Z_2, Z_3	Z_1	Z_1, Z_5	Z_2, Z_6	Z_2	Z_2, Z_4, Z_5	Z_2	Z_2, Z_4, Z_5
002	Z_2, Z_4, Z_5	Z_1, Z_3	Z_1, Z_5	Z_4, Z_5, Z_6	Z_1, Z_2	Z_3	Z_2, Z_3, Z_6	Z_1, Z_2	Z_4, Z_6
003	Z_1	Z_2	Z_1, Z_4, Z_5	Z_1, Z_5	Z_3, Z_5	Z_2, Z_4, Z_5	Z_2, Z_3	Z_2, Z_3, Z_5	Z_2, Z_3, Z_6
004	Z_3, Z_5	Z_2, Z_3	Z_1	Z_1, Z_5	Z_3	Z_3	Z_2, Z_4, Z_5	Z_1, Z_3	Z_3
005	Z_1, Z_5	Z_3	Z_1, Z_2	Z_6	Z_3, Z_5	Z_3, Z_5	Z_1, Z_2, Z_6	Z_3	Z_4, Z_6
006	Z_4	Z_2, Z_3	Z_5	Z_2, Z_3, Z_4	Z_1, Z_3	Z_2, Z_3	Z_3	Z_3	Z_2, Z_5
007	Z_2, Z_3, Z_5	Z_2, Z_3, Z_4	Z_1, Z_2, Z_3	Z_6	Z_1, Z_5	Z_2, Z_3	Z_3, Z_4	Z_4	Z_3, Z_4
…	…	…	…						

2.5.2 Fuzzy associate rule based for KANSEI evaluation

In the current work, the relation between the certain adjective and a certain form feature is described. Consider Z_k to A_{ij} which can be presented as $Z_k \Rightarrow A_{ij}$, thus a fuzzy presentation was constructed for the relationship between KANSEI adjectives and the forms' features, support and confidence of the rules were given as follows.

Let the score of the evaluation of KANSEI Z_k on the form feature A_{ij} to be $score_{ij}^{(k)} k = 1, 2, \ldots, 6$, $i = 1, 2, \ldots, 12$, $j = 1, 2, \ldots, 9$, where $N\_score_{ij}^{(k)}$ is normalization of $score_{ij}^{(k)}$ which was calculated as:

$$N\_score_{ij}^{(k)} = \frac{score_{ij}^{(k)} - 1}{6} \tag{29}$$

Let $v_{ij}^{(k)}$ be the membership of form feature A_{ij} about KANSEI Z_k, which is given by:

$$v_{ij}^{(k)} = \frac{\sum N\_score_{ij}^{(k)}}{Count(D)} \tag{30}$$

A fuzzy presentation on the form's feature A_{ij} for KANSEI Z_k is generated as follow:

$$\left\{ \frac{v_{ij}^{(1)}}{Z_1} + \frac{v_{ij}^{(2)}}{Z_2} + \frac{v_{ij}^{(3)}}{Z_3} + \frac{v_{ij}^{(4)}}{Z_4} + \frac{v_{ij}^{(5)}}{Z_5} + \frac{v_{ij}^{(6)}}{Z_6} \right\} \tag{31}$$

The $Z_k \Rightarrow A_{ij}$ rules were generated continuously, thus, the support can be calculated using:

$$Support\,(Z_k \Rightarrow A_{ij}) = v_{ij}^{(k)} \tag{32}$$

In addition, the confidence is given by:

$$Confidence\,(Z_k \Rightarrow A_{ij}) = \frac{v_{ij}^{(k)}}{\sum_{l=1}^{6} v_{ij}^{l}} \tag{33}$$

3 Results and discussion

The proposed approach is applied in the dataset for the shoe forms shown in Fig. 5, where the extracted features are illustrated in Fig. 6. In order to evaluate the KANSEI based on the fuzzy associate rule, the candidate set C_1 is acquired in Table 2.

In the present work,$12 \times 9 \times 6 = 648$ rules were exist, in addition set the threshold of support to be 0.3, and the confidence be 0.55, all support less than 0.3 or confidence less than 0.55 were deleted. Thus, the obtained value for L_1 is 277, and using the proposed algorithm in Sect. 2.4.1 to obtain 2-rank candidate set C_2, the updated support and confidence was calculated as follows:

$$Support\,(Z_k, Z_p \Rightarrow A_{ij}) = \min\{Support\,(Z_i \Rightarrow A_{ij}) | i = k, p\} \tag{34}$$

$$Confidence\,(Z_k, Z_p \Rightarrow A_{ij}) = \min\{Confidence\,(Z_i \Rightarrow A_{ij}) | i = k, p\} \tag{35}$$

Some KANSEI adjectives were deleted for lower support and confidence in 1-rank high frequent set L_1, and the $Support = 0.45$ and $Confident = 0.65$ were utilized to obtain 109 2-rank high frequent set L_2 in Table 3 and relative products list and form features in Fig. 7.

Table 2 From candidate item-set C_1 to frequent item-set L_1

Form feature	Rules	Support	Confidence	Operation
A_{11}	$Z_1 \Rightarrow A_{11}$	0.31	0.47	*Delete*
	$Z_2 \Rightarrow A_{11}$	0.52	0.68	Reserve
	$Z_3 \Rightarrow A_{11}$	0.65	0.81	Reserve
	$Z_4 \Rightarrow A_{11}$	0.76	0.76	Reserve
	$Z_5 \Rightarrow A_{11}$	0.28	0.47	*Delete*
	$Z_6 \Rightarrow A_{11}$	0.22	0.55	*Delete*
A_{12}	$Z_1 \Rightarrow A_{12}$	0.52	0.81	Reserve
	$Z_2 \Rightarrow A_{12}$	0.89	0.68	Reserve
	$Z_3 \Rightarrow A_{12}$	0.45	0.66	Reserve
	$Z_4 \Rightarrow A_{12}$	0.27	0.57	*Delete*
	$Z_5 \Rightarrow A_{12}$	0.39	0.53	*Delete*
	$Z_6 \Rightarrow A_{12}$	0.50	0.68	Reserve
…	…			
$A_{12,9}$	$Z_1 \Rightarrow A_{12,9}$	0.52	0.72	Reserve
	$Z_2 \Rightarrow A_{12,9}$	0.44	0.44	*Delete*
	$Z_3 \Rightarrow A_{12,9}$	0.26	0.18	*Delete*
	$Z_4 \Rightarrow A_{12,9}$	0.67	0.76	Reserve
	$Z_5 \Rightarrow A_{12,9}$	0.24	0.43	*Delete*
	$Z_6 \Rightarrow A_{12,9}$	0.52	0.77	Reserve

Table 3 From 2-candidate item-set C_2 to frequent item-set L_2

Form features	Rules	Support	Confidence	Operation
A_{11}	$Z_2, Z_3 \Rightarrow A_{11}$	0.55	0.78	Reserve
	$Z_2, Z_4 \Rightarrow A_{11}$	0.52	0.57	*Delete*
	$Z_3, Z_4 \Rightarrow A_{11}$	0.70	0.72	Reserve
A_{12}	$Z_1, Z_2 \Rightarrow A_{12}$	0.52	0.67	Reserve
	$Z_1, Z_3 \Rightarrow A_{12}$	0.55	0.71	Reserve
	$Z_2, Z_3 \Rightarrow A_{12}$	0.44	0.70	*Delete*
…	…			
$A_{12,9}$	$Z_1, Z_3 \Rightarrow A_{12,9}$	0.54	0.71	Reserve
	$Z_1, Z_5 \Rightarrow A_{12,9}$	0.41	0.74	*Delete*

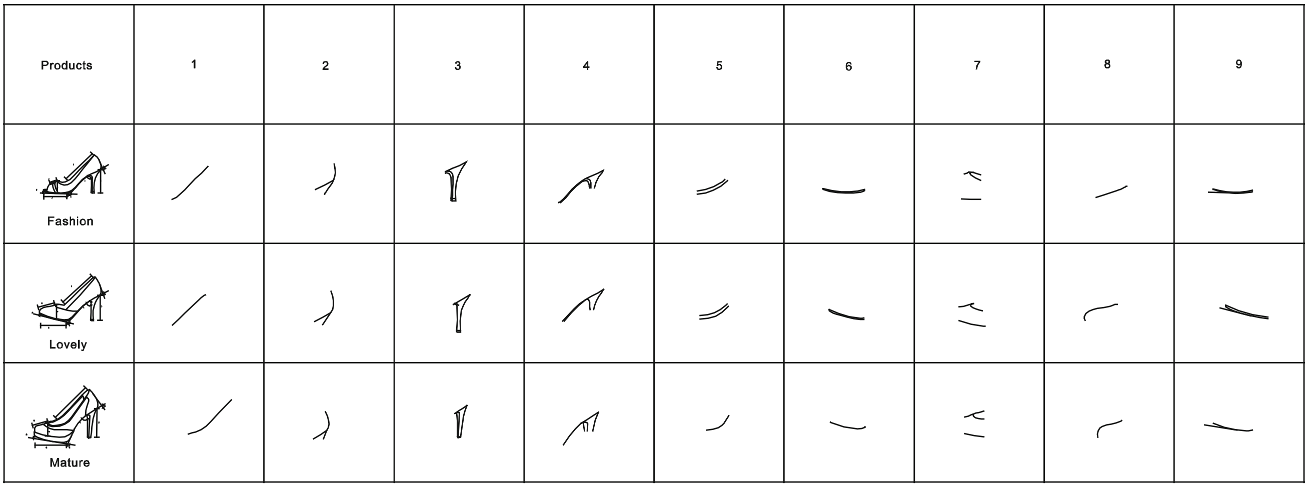

Fig. 7 The products based on 2-rank high frequent rules

Table 4 The 2-frequent item-set

Rule	Support	Confidence	Rule	Support	Confidence
$Z_2, Z_3 \Rightarrow A_{11}$	0.50	0.69	$Z_4, Z_6 \Rightarrow A_{11}$	0.55	0.76
$Z_2, Z_4 \Rightarrow A_{11}$	0.51	0.60	$Z_3, Z_4 \Rightarrow A_{11}$	0.56	0.80
$Z_3, Z_4 \Rightarrow A_{11}$	0.74	0.78	$Z_1, Z_2 \Rightarrow A_{12}$	0.58	0.69
$Z_1, Z_2 \Rightarrow A_{12}$	0.51	0.69	$Z_2, Z_3 \Rightarrow A_{12}$	0.56	0.83
$Z_1, Z_3 \Rightarrow A_{12}$	0.52	0.65	$Z_2, Z_4 \Rightarrow A_{12}$	0.56	0.76
$Z_3, Z_4 \Rightarrow A_{13}$	0.53	0.77	$Z_1, Z_6 \Rightarrow A_{13}$	0.59	0.68
$Z_1, Z_4 \Rightarrow A_{14}$	0.55	0.67	$Z_3, Z_4 \Rightarrow A_{14}$	0.66	0.78
$Z_3, Z_6 \Rightarrow A_{21}$	0.64	0.67	$Z_3, Z_4 \Rightarrow A_{21}$	0.67	0.87
$Z_1, Z_3 \Rightarrow A_{22}$	0.54	0.66	$Z_2, Z_3 \Rightarrow A_{22}$	0.79	0.86
$Z_2, Z_6 \Rightarrow A_{22}$	0.54	0.78	…	…	…
…	…	…	…	…	…
$Z_1, Z_4 \Rightarrow A_{15,9}$	0.52	0.79	$Z_2, Z_3 \Rightarrow A_{15,9}$	0.65	0.82

In addition, the back propagation neural networks training results are obtained using the above high-frequency rule set to train the BPNN in order to predict the high-frequency rule in the year 2016 by using the survey data of 2014 and 2015. The neuron input index system is given, which is defined based on the support in Eq. (32), the confidence in Eq. (33), and the improvement on rule $Z_i, Z_k \Rightarrow A_{jk}$ which is defined as:

$$Improvement = \frac{Confidence}{Count\left(\sum_{Z_i, Z_p \in A_{jk}} A_{jk}\right)} \tag{36}$$

In addition, the recentness defined as to rule that has recently generated a date from the present and the frequency of this rule, are also included. The data from years 2014 and 2015 are imported to generate high frequent set L_2 as demonstrated in Table 4.

Let the two periods of 2-rank rule set's support and confidence be the BPNN training data set, they will be trained separately, i.e., support value training and confidence value training, while the support and confidence time series are predicted, the high frequent rule set is reserved by using the same threshold for the past and all the rules with KANSEI adjective Z_1 from 2 rank high frequent item-set from 2013, 2014 and 2015 as targets for the same rules. The BPNN is used to predict high frequent set of form features through support and confidence threshold filter. The nonlinear autoregressive is used in the current work to predict the support and confidence of the same rule with the

Table 5 Values of BPNN training system

	Value		Value
Epoch	372	Mu	0.005
Time	6	Effective # Param	72
Performance	249e+5	Sum Square Param	93.2
Gradient	4.68e+5	Validation checks	0

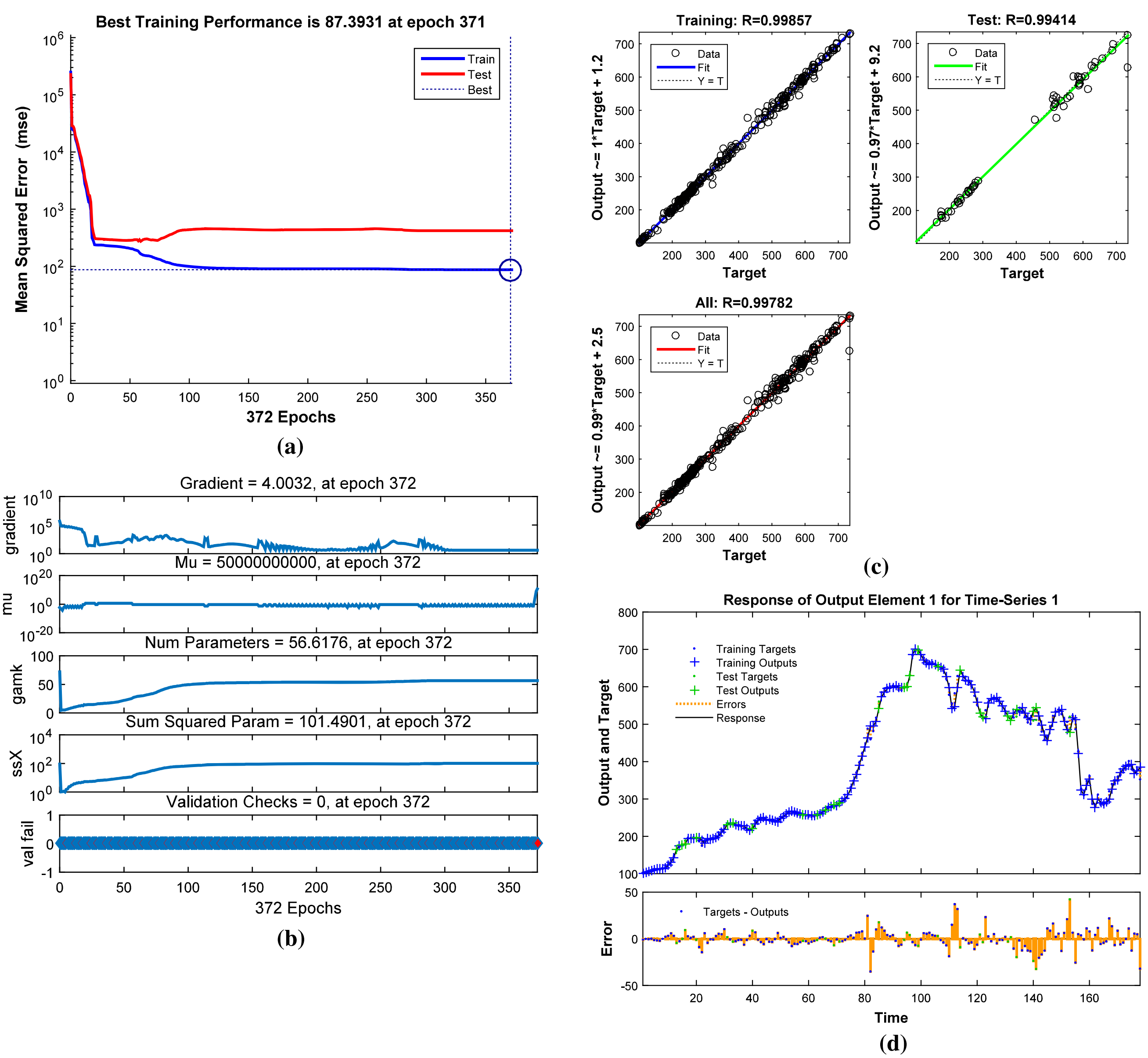

Fig. 8 The BPNN training results for time series support and confidence with all rules by Z_1 where **a** is MSE performance for epoch; **b** is training state; **c** is regressive process and **d** is time series response

same KANSEI adjectives. The number of hidden neuron is 10, number of delays is 2, training algorithm is Bayesian regularization, set division randomness and mean square error (MSE) as performance, and 1000 epochs. Table 5 and Fig. 8 illustrate the obtained results.

Finally, the neural network predictive value of the operation is kept or is deleted to obtain the high-frequency rule set in 2016 as shown in Table 6 and the relative products and form features listed in Fig. 9.

Normally, typical BPNN training algorithms include Bayesian regularization (BR), Levenberg–Marquardt (LM) and Scaled Conjugate Gradient (SCG). The BR is training function of a network that minimizes a combination of MSE and then determines the correct combination to produce a superior network through updating the weight and bias values. Levenberg–Marquardt algorithm is one of the

Table 6 2-Frequent item-set by predicting

Rules
$Z_1, Z_4 \Rightarrow A_{13} Z_3, Z_4 \Rightarrow A_{14} Z_3, Z_5 \Rightarrow A_{33} Z_2, Z_3 \Rightarrow A_{21}$
$Z_2, Z_5 \Rightarrow A_{24} Z_1, Z_3 \Rightarrow A_{41} Z_1, Z_6 \Rightarrow A_{42} Z_2, Z_5 \Rightarrow A_{45}$
$Z_1, Z_2 \Rightarrow A_{53} Z_2, Z_3 \Rightarrow A_{64} Z_1, Z_2 \Rightarrow A_{72} Z_1, Z_2 \Rightarrow A_{83}$
$Z_4, Z_6 \Rightarrow A_{85} Z_3, Z_5 \Rightarrow A_{97} Z_5, Z_6 \Rightarrow A_{10,2} Z_2, Z_6 \Rightarrow A_{10,4}$
$Z_3, Z_5 \Rightarrow A_{11,2} Z_5, Z_6 \Rightarrow A_{11,8} Z_1, Z_4 \Rightarrow A_{12,7} Z_2, Z_4 \Rightarrow A_{12,8}$

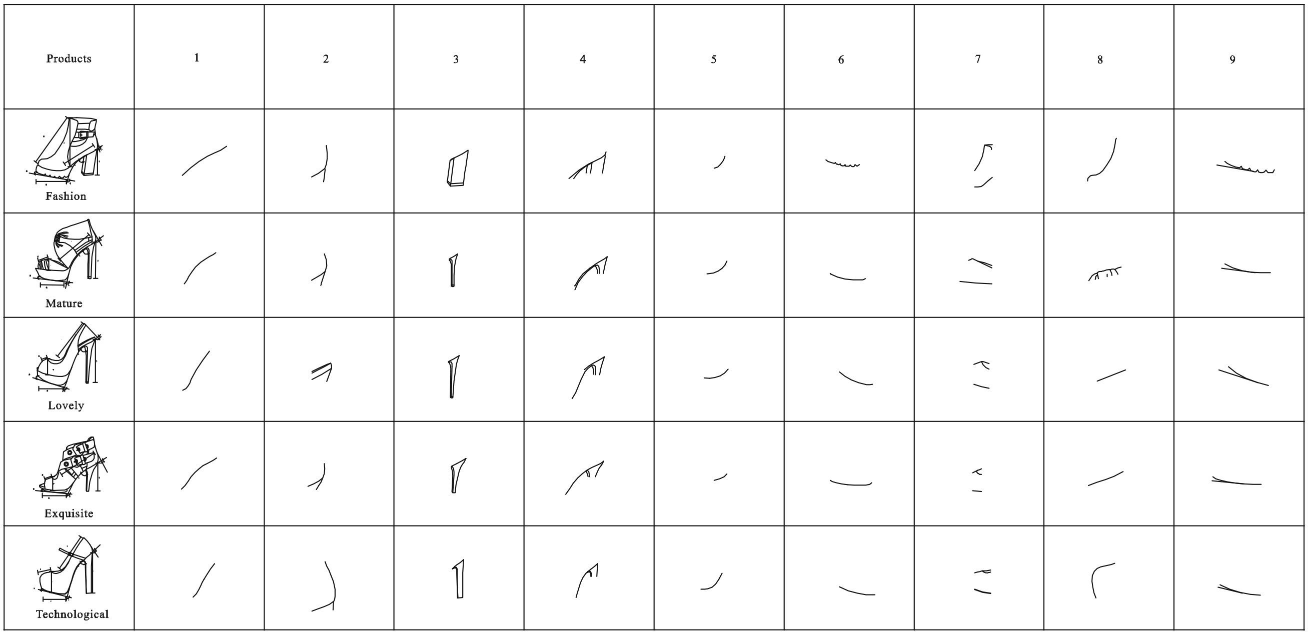

Fig. 9 Time series support and confidence-based BPNN prediction for form features and products with KANSEI adjectives

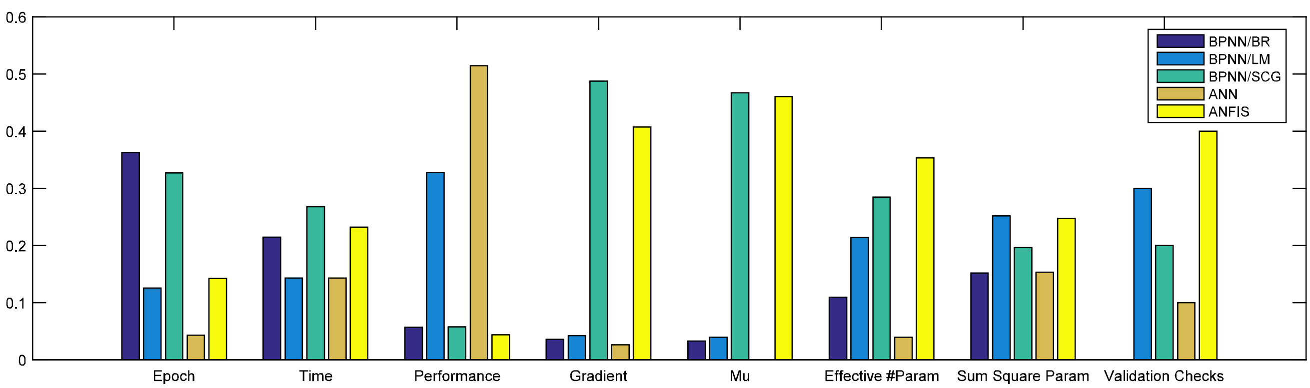

Fig. 10 Comparing with several data pre-processing of BPNN other neural network algorithms

optimization algorithms to search for the minimum of the parameters of the function values. In addition, the most widely used nonlinear least squares algorithm is the gradient of the largest (small) value of the algorithm. It has the advantages of both the gradient method and the Newton method. When the time is very small, the step size is equal to the Newton step size, while with long time, the step size is equal to the gradient descent method. In addition, the conjugate gradient method is an algorithm for the numerical solution of particular systems of linear equations. It can also be used to solve unconstrained optimization problems such as energy minimization. The ANN is a kind of artificial neural network that is based on Takagi–Sugeno fuzzy inference system for normal training. In the current work, all indices for BPNN were normalized with Bayesian regularization noted BPNN/BR, Levenberg–Marquardt as BPNN/LM, scaled conjugate gradient as BPNN/SCG, artificial neural networks.

In the current work, the ANN and the adaptive neuron fuzzy inference system (ANFIS) for the same dataset of confidence and support value of KANSEI Z_1 are applied. Figure 10 demonstrates the comparison of the proposed approach (BPNN/BR) with several neural network algorithms.

The proposed BPNN/BR approach achieves superior performance to the other algorithms as illustrated in Fig. 9 in terms of the performance, gradient, Mu, Effective number of parameter, and the sum square parameter in KANSEI support and confidence time series prediction. In the current work, the proposed variable rough set reduction algorithm also scaled the complexity of the rule set, and using Bayesian regularization, BPNN shows the good performance in KANSEI evaluation.

However, some improvement for the existing work in the future work can be performed, namely: (i) the sample of the products' evaluation data which acquired from a questionnaire system needs to be improved, and big data technology or transaction database would be applied and (ii) the neural network training system requires validation in some special issues such as the stability of the inference system, while the input data include isolated point, and fuzzy presentation method also needs to be improved in membership definition, and (iii) for variable precision rough set, the key issue on the research is to find or give the upper approximation set, lower approximation set and boundary set, so the more test need to be assigned for getting more precision value for the boundary.

4 Conclusions

The KANSEI evaluation is concerning the scaling system. Previously, linear optimization was utilized. In the present work, a new knowledge presentation method with KANSEI evaluation by using variable precision rough set was applied. The rule-based presentation was deducted, where the form features as attributions were also reduced for the next inference system. The nonlinear inference system BPNN with Bayesian regularization was proposed to predict high frequent set of the forms' features with certain KANSEI adjectives. In case study, rules were reduced and high-frequency rule set was acquired by using the proposed algorithms in this research; time series support and confidence-based BPNN prediction for form features and products with KANSEI adjectives were also produced finally. The evaluation comparative results established that the proposed method is effective compared to other neural network training systems. Although the neural network system or other nonlinear inference system were applied in shoes' industry, but it is only for selecting some shoes' materials, while this paper incorporating KANSEI engineering method to shoe industry. As result, the effectiveness and feasibility of the proposed methods in this paper is reasonable and has clarity.

Acknowledgements This work was supported by Zhejiang Provincial Natural Science Fund of China under No. LY17F030014.

Compliance with ethical standards

Conflict of interest This work has no conflict of interest.

References

1. Baek S, Hwang M, Chung H, Kim P (2008) Kansei factor space classified by information for kansei image modeling. Appl Math Comput 205:874–882. Special issue on advanced intelligent computing theory and methodology in applied mathematics and computation
2. Barton J, Lees A (1996) Comparison of shoe insole materials by neural network analysis. Med Biol Eng Comput 34:453–459
3. Robson B, Boray S (2016) Data-mining to build a knowledge representation store for clinical decision support. Studies on curation and validation based on machine performance in multiple choice medical licensing examinations. Comput Biol Med 73:71–93
4. Keshavamurthy BN, Khan AM, Toshniwal D (2013) Privacy preserving association rule mining over distributed databases using genetic algorithm. Neural Comput Appl 22:351–364
5. Khadse CB, Chaudhari MA, Borghate VB (2016) Conjugate gradient back-propagation based artificial neural network for real time power quality assessment. Int J Electr Power Energy Syst 82:197–206
6. Wang C, Wu F, Shi Z et al (2016) Indoor positioning technique by combining RFID and particle swarm optimization-based back propagation neural network. Optik 127(17):6839–6849
7. Chen D, Zhang X, Li W (2015) On measurements of covering rough sets based on granules and evidence theory. Inf Sci 317:329–348
8. Chen F-H, Chi D-J, Wang Y-C (2015) Detecting biotechnology industry's earnings management using Bayesian network, principal component analysis, back propagation neural network, and decision tree. Econ Model 46:1–10
9. Chen M-C, Hsu C-L, Chang K-C, Chou M-C (2015) Applying Kansei engineering to design logistics services c a case of home delivery service. Int J Ind Ergon 48:46–59
10. Chu C-H, Tsai Y-T, Wang CC, Kwok T-H (2010) Exemplar-based statistical model for semantic parametric design of human body. Comput Ind, 61:541–549. Soft Products Development
11. Czibula G, Czibula IG, Sîrbu A-M, Mircea I-G (2015) A novel approach to adaptive relational association rule mining. Appl Soft Comput 36:519–533
12. Dahal K, Almejalli K, Hossain MA, Chen W (2015) Ga-based learning for rule identification in fuzzy neural networks. Appl Soft Comput 35:605–617
13. Farid DM, Zhang L, Rahman CM, Hossain M, Strachan R (2014) Hybrid decision tree and naïve Bayes classifiers for multi-class classification tasks. Expert Syst Appl 41:1937–1946
14. Kargarfard F, Sami A, Ebrahimie E (2015) Knowledge discovery and sequence-based prediction of pandemic influenza using an integrated classification and association rule mining (CBA) algorithm. J Biomed Inform 57:181–188
15. Fung K, Kwong C, Siu K, Yu K (2012) A multi-objective genetic algorithm approach to rule mining for affective product design. Expert Syst Appl 39:7411–7419
16. Fuqian Shi JX, Sun S (2012) Employing rough sets and association rule mining in Kansei knowledge extraction. Inf Sci 196:118–128
17. He T, Cao L, Balas VE, McCauley P, Shi F (2016) Curvature manipulation of the spectrum of valence–arousal-related fMRI dataset using Gaussian-shaped fast Fourier transform and its application to fuzzy KANSEI adjectives modeling. Neurocomputing 174:1049–1059
18. Hong-Bin Yan V-NH, Nakamori Y (2012) A group non additive multi-attribute consumer-oriented Kansei evaluation model with an application to traditional crafts. Ann Oper Res 195:325–354

19. Huang KY, Li I-H (2016) A multi-attribute decision-making model for the robust classification of multiple inputs and outputs datasets with uncertainty. Appl Soft Comput 38:176–189
20. Huang Y, Chen C-H, Khoo LP (2012) Kansei clustering for emotional design using a combined design structure matrix. Int J Ind Ergon 42:416–427
21. Huang Y, Chen C-H, Wang I-HC, Khoo LP (2014) A product configuration analysis method for emotional design using a personal construct theory. Int J Ind Ergon 44:120–130
22. Imai M, Imai Y, Hattori T (2013) Collaborative design and its evaluation through Kansei engineering approach. Artif Life Robot 18:233–240
23. Jiang H, Kwong C, Law M, Ip W (2013) Development of customer satisfaction models for affective design using rough set and {ANFIS} approaches. Proc Comput Sci 22:104–112. In: 17th international conference in knowledge based and intelligent information and engineering systems—KES2013
24. Jiang L, Wang S, Li C, Zhang L (2015) Structure extended multinomial naive Bayes. Inf Sci 329:346–356
25. Cheng Jun, Wang Xin, Si Tingting et al (2016) Ignition temperature and activation energy of power coal blends predicted with back-propagation neural network models. Fuel 173:230–238
26. Koc L, Mazzuchi TA, Sarkani S (2012) A network intrusion detection system based on a hidden naïve Bayes multiclass classifier. Expert Syst Appl 39:13492–13500
27. Lee C-H (2015) A gradient approach for value weighted classification learning in naive Bayes. Knowl-Based Syst 85:71–79
28. Leema N Khanna, Nehemiah H, Kannan A (2016) Neural network classifier optimization using differential evolution with global information and back propagation algorithm for clinical datasets. Appl Soft Comput. doi:10.1016/j.asoc.2016.08.001
29. Li W, Xu W (2015) Double-quantitative decision-theoretic rough set. Inf Sci 316:54–67. Nature-Inspired Algorithms for Large Scale Global Optimization
30. Zhou Ligang, Dong Lu, Fujita Hamido (2015) The performance of corporate financial distress prediction models with features selection guided by domain knowledge and data mining approaches. Knowl-Based Syst 85:52–61
31. Lin M-C, Lin Y-H, Lin C-C, Chen M-S, Hung Y-C (2015) An integrated neuro-genetic approach incorporating the Taguchi method for product design. Adv Eng Inform 29:47–58
32. Lin S-H, Huang C-C, Che Z-X (2015) Rule induction for hierarchical attributes using a rough set for the selection of a green fleet. Appl Soft Comput 37:456–466
33. Lu W, Petiot J-F (2014) Affective design of products using an audio-based protocol: application to eyeglass frame. Int J Ind Ergon 44:383–394
34. Luo S-J, Fu Y-T, Korvenmaa P (2012) A preliminary study of perceptual matching for the evaluation of beverage bottle design. Int J Ind Ergon 42:219–232
35. Luximon Y, Cong Y, Luximon A, Zhang M (2015) Effects of heel base size, walking speed, and slope angle on center of pressure trajectory and plantar pressure when wearing high-heeled shoes. Hum Mov Sci 41:307–319
36. del Rosario Martínez-Blanco M, Ornelas-Vargas G, Solís-Sánchez LO et al (2016) A comparison of back propagation and generalized regression neural networks performance in neutron spectrometry. Appl Radiat Isot 117:20–26
37. Meng-Dar Shieh Y-EY, Huang C-L (2016) Eliciting design knowledge from affective responses using rough sets and Kansei engineering system. J Ambient Intell Human Comput 7(1):107–120
38. Nagamachi M (1995) Kansei engineering: a new ergonomic consumer oriented technology for product development. Int J Ind Ergon 15:3–11. Kansei engineering: an ergonomic technology for product development
39. Nagamachi M (2002) Kansei engineering as a powerful consumer-oriented technology for product development. Appl Ergon 33:289–294. Fundamental reviews in applied ergonomics 2002
40. Nguyen D, Nguyen LT, Vo B, Hong T-P (2015) A novel method for constrained class association rule mining. Inf Sci 320:107–125
41. Nguyen LT, Nguyen NT (2015) An improved algorithm for mining class association rules using the difference of obidsets. Expert Syst Appl 42:4361–4369
42. Nishiwaki T, Nonogawa M (2015) Application of topological optimization technique to running shoe designing. Proc Eng 112:314–319. In: 'The impact of technology on sport VI' 7th Asia-Pacific congress on sports technology, APCST2015
43. Nithya NS, Duraiswamy K (2014) Gain ratio based fuzzy weighted association rule mining classifier for medical diagnostic interface. Sadhana 39:39–52
44. Ogunde A, Folorunso O, Sodiya A (2015) A partition enhanced mining algorithm for distributed association rule mining systems. Egypt Inform J 16(3):297–307
45. Pan D, Liu D, Zhou J, Zhang G (2015) Anomaly detection for satellite power subsystem with associated rules based on kernel principal component analysis. Microelectron Reliab 55:2082–2086. In: Proceedings of the 26th European symposium on reliability of electron devices, failure physics and analysis si: proceedings of ESREF 2015
46. Kaur Parneet, Singh Manpreet, Josan Gurpreet Singh (2015) Classification and prediction based data mining algorithms to predict slow learners in education sector. Proc Comput Sci 57:500–508
47. Sagara T, Hagiwara M (2014) Natural language neural network and its application to question-answering system. Neurocomputing 142:201–208. SI: computational intelligence techniques for new product development
48. Sahoo J, Das AK, Goswami A (2015) An efficient approach for mining association rules from high utility itemsets. Expert Syst Appl 42:5754–5778
49. Shieh M-D, Yeh Y-E (2013) Developing a design support system for the exterior form of running shoes using partial least squares and neural networks. Comput Ind Eng 65:704–718
50. Smith S, Fu S-H (2011) The relationships between automobile head-up display presentation images and drivers Kansei. Displays 32:58–68
51. Son NN, Anh HPH (2015) Adaptive displacement online control of shape memory alloys actuator based on neural networks and hybrid differential evolution algorithm. Neurocomputing 166:464–474
52. Stepnicka M, Burda M, Stepnickova L (2015) Fuzzy rule base ensemble generated from data by linguistic associations mining. Fuzzy Sets Syst 285(C):140–161
53. Suresh A, Harish K, Radhika N (2015) Particle swarm optimization over back propagation neural network for length of stay prediction. Proc Comput Sci 46:268–275. In: Proceedings of ICICT 2014, 3–5 December 2014 at Bolgatty Palace and Island Resort, Kochi, India
54. Tang C, Fung K, Lee EW, Ho G, Siu KW, Mou W (2013) Product form design using customer perception evaluation by a combined super ellipse fitting and ANN approach. Adv Eng Inform 27:386–394
55. Tzu-Hsuan Huang KS, Yokota M (2012) Toward artificial Kansei based on mental image directed semantic theory. Artif Life Robot 17:186–190
56. Ushada M, Murase H (2008) Pattern extraction from human preference reasoning using conditional probability co-occurrences matrix of texture analysis. Eng Agric Environ Food 1:45–50

57. Ushada M, Okayama T, Suyantohadi A, Khuriyati N, Murase H (2015) Daily worker evaluation model for SME-scale food production system using Kansei engineering and artificial neural network. Agric Agric Sci Proc 3:84–88. In: International conference on agro-industry (IcoA): Sustainable and Competitive Agro-industry for Human Welfare Yogyakarta-INDONESIA 2014
58. Wang K-C (2011) A hybrid Kansei engineering design expert system based on grey system theory and support vector regression. Expert Syst Appl 38:8738–8750
59. Wang L, Zeng Y, Chen T (2015) Back propagation neural network with adaptive differential evolution algorithm for time series forecasting. Expert Syst Appl 42:855–863
60. Ziarko Wojciech (1993) Variable precision rough set model. J Comput Syst Sci 46(1):39–59
61. Wu J, Pan S, Zhu X, Cai Z, Zhang P, Zhang C (2015) Self-adaptive attribute weighting for naïve Bayes classification. Expert Syst Appl 42:1487–1502
62. Yao Y, She Y (2016) Rough set models in multi granulation spaces. Inf Sci 327:40–56
63. Zhai L-Y, Khoo L-P, Zhong Z-W (2009) A dominance-based rough set approach to Kansei engineering in product development. Expert Syst Appl 36:393–402
64. Zhai L-Y, Khoo L-P, Zhong Z-W (2009) A rough set based decision support approach to improving consumer affective satisfaction in product design. Int J Ind Ergon 39:295–302
65. Zhang H, Zhou J, Miao D, Gao C (2012) Bayesian rough set model: a further investigation. Int J Approximate Reasoning 53:541–557
66. Zhang X, Hao S, Xu C, Qian X, Wang M, Jiang J (2015) Image classification based on low-rank matrix recovery and naive Bayes collaborative representation. Neurocomputing 169:110–118