Appl Intell (2011) 35:75–88
DOI 10.1007/s10489-009-0204-9

Building a qualitative recruitment system via SVM with MCDM approach

Yung-Ming Li · Cheng-Yang Lai · Chien-Pang Kao

Published online: 29 January 2010

Abstract Advances in information technology have led to behavioral changes in people and submission of curriculum vitae (CV) via the Internet has become an often-seen phenomenon. Without any technological support for the filtering process, recruitment can be difficult. In this research, a method combining five-factor personality inventory, support vector machine (SVM), and multi-criteria decision-making (MCDM) method was proposed to improve the quality of recruiting appropriate candidates. The online questionnaire personality testing developed by the International Personality Item Pool (IPIP) was utilized to identify the personal traits of candidates and both SVM and MCDM were employed to predict and support the decision of personnel choice. SVM was utilized to predict the fitness of candidates, while MCDM was employed to estimate the performance for a job placement. The results show the proposed system provides a qualified matching according to the results collected from enterprise managers.

Keywords Support vector machine · TOPSIS · Five-factor personality inventory · Candidate recruiting · Personality trait

Y.-M. Li (✉) · C.-Y. Lai · C.-P. Kao
Institute of Information Management, National Chiao Tung University, Hsinchu, Taiwan
e-mail: yml@mail.nctu.edu.tw

C.-Y. Lai
e-mail: cylai.iim96g@g2.nctu.edu.tw

C.-P. Kao
e-mail: jovi.iim95g@nctu.edu.tw

1 Introduction

Over the years, personality trait has already been the best checkered reputation as a predictor of work performance. In 2005, over 40 percent of CEOs stated that the most important operational challenge was to find, hire, and retain a qualified employee [12]. About 66 percent of CEOs listed high-quality employees as the most important factor contributing to the growth of their company [12]. Ultimately, improving the quality of recruitment would allocate the right personnel to the right position, which consequently enables the employees to achieve greater work performances and significantly reduce the employee training cost. Therefore, a highly qualitative process of employee hiring and allocation would effectively increase the core competitiveness of the firms and it is also very beneficial to the firms for facing today's dynamic global marketplace.

Companies are made up of people. In order to operate a company successfully and smoothly, managers have to hire suitable employees with different expertise. The cooperation of employees and managers allows the enterprises to achieve the business goal and vision. In order to reach this goal, it is necessary for enterprises to achieve the effective performance of employees. Effectiveness in work performance begins with qualitative employee selection. Therefore, personnel directors are constantly faced with the problem of choosing the appropriate personnel from a mutually exclusive set of personnel pool. However, it is very time-consuming and laborious to identify candidates with the right potential traits. Consequently, an improvement in the quality of hiring can be a great work performance booster for the organization and assure the firm of its continual growth.

The match among an employee's knowledge, skills, and abilities (KSA), and other more abstract characteristics, such as personality and value orientations, has been used as a major recruiting criterion in the past three decades. Personality characteristics and the corresponding measures have increasingly been used by human resource professionals and managers to evaluate the suitability of job applicants. Human choice behavior could be seen as a process of identifying the significant difference between alternatives [23]. Personality characteristics can be utilized to screen personnel differentiation. Over the years, researchers have acknowledged and documented the fact that personality could be a good predictor of work outcomes in a wide variety of jobs ranging from skilled, semiskilled to executives. Specifically, the taxonomy of personality characteristics, such as Five-Factor Model (FFM) or the Big Five, has motivated a series of meta-analytic studies. While these studies provided a much more optimistic view of the ability of personality measures to predict job performance, we intend to propose an expert system to facilitate personnel recruitment for human resource development.

Most decision-making problems or choice problems faced in the real world fall into the multi-attribute evaluation category. For the personnel choice problem, the factors to be considered usually include the personality factors, traits and skills that can be seen as the attributes of a person. In addition, recruiting a qualified person or promoting a suitable employee to a new position could be portrayed as a decision-making process. In the paper, exploiting the learning ability of the Support Vector Machine (SVM) and the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) method of multi-criteria decision-making (MCDM), a qualitative personnel recruitment system was proposed. The proposed framework could be used to discover the suitable work/department for different types of personnel. Also, it could be used to estimate the possible work performance of personnel, if he/she is located to some specific position.

Personality is one of the most important influential factors when people are facing work pressure. That is, personality significantly affects human behaviors and attitudes in both daily life and work. In this study, personality trait is utilized as a significant element for building the departments' human resource personality model. This model aims to predict the fitness of a job position with either current employees or new job applicants. According to the results of personality match and performance prediction, the proposed system can give managers valuable suggestions of personnel recruitment. Managers and the high-level executives can understand and discover the potential of existing employees easily from the proposed model when they need to carry out job adjustments within the organization or select the most appropriate candidates from new applicants. It is especially useful when managers decide to promote a qualified employee to a new job or a new position, or to choose the suitable employees to form a new team for a specific project.

The remaining part of this paper is organized as follows. Section 2 introduces related works. In Sect. 3 we detail the whole framework combining FFM personality trait, SVM and MCDM. Section 4 describes the practical experiments, and results and evaluations. Section 5 draws the conclusions of this study and offers suggestions for further research.

2 Related work

2.1 Personality trait

Personality is often defined as a group of characteristics that structure one's reactions to oneself. Dunn et al. [13] have shown that during the hiring decision, managers weight individual personality characteristics as if they were as important as general mental ability. Furthermore, research shows that personality contributes incremental validity in the prediction of job performance above and beyond that accounted for by other predictors, including general mental ability and biodata [28, 30, 32]. Barrick and Mount [2] summarized the role of personality at work in seven divergent research streams to demonstrate that personality matters because it predicts and explains the behaviors at work. Thus, in order to resolve the problem of the match between personality of employee and jobs, we focus on the personality trait of candidates. In addition to conventional "Curriculum Vitae-Job (CV-Job)" matching, we discuss its role in personnel selection and its correction with job performance.

Pre-recruitment activities are increasingly emphasized as the first step in the hiring process. Beagrie [3] has estimated that two thirds of medium to large organizations use some kind of psychological testing, including aptitude as well as personality in-job applicant screening. Another survey indicates that all top 100 companies in Great Britain reported using personality tests as part of their hiring procedure [12, 31]. One of the most prevalent reasons for using personality testing is its contribution to improve the fitness between personality traits of job-seeker and work position. That will further increase the work satisfaction rate and reduce the turnover rate. The personality should be examined before making a hiring decision. It appears that personality testing has been increasingly used as a component of the personnel selection process.

Tzeng et al. [36] used SVM to predict the turnover rate of nurses. Their model used working motivation, job satisfaction, and stress levels as classifiers to predict the intention of job withdrawal of nurses. However, it did not consider the personality of the nurses. In other words, it neglects the personality effect of nurses' suitability for their job or the work of health care in this study. Hong et al. [16] proposed another method for predicting employees' stay or leave in an organization according to job performance. In these works, the SVM classifiers only utilize the job performances of staff as the basis for categorization. The possibility of intention to leave an organization of a worker is judged according to working performance alone. The most basic characteristic of people was disregarded. That is, it did not take reasons for the lower work performance into account-whether the personality of employees is unsuitable for their jobs.

2.2 Personality assessment and five factor model

Seminal work in clarifying the five dimensions has been proposed by Costa and McCrae [10]. They developed a model of personality known as the "Five Factor Model of Personality" namely FFM or Big Five trait. FFM reveals small to nonexistent mean score differences between the racial or ethnic groups [17, 18, 29]. Barrick and Mount [2] argued that it is of enormous significance because most organizations are keenly interested in hiring a more diverse workforce. Goodstein and Lanyon also credit the FFM for providing a universally accepted set of dimensions for describing human behaviors at work [14].

There are five taxonomies of personality characteristics identified in this personality assessment model, such as Extraversion, Agreeableness, Emotional Stability, Conscientiousness, and Openness to Experience. They provide a global description of basic personality traits in a more general construct. After completing the personality assessment through the IPIP-NEO questionnaire, the testing results with 30 scores are classified into the five factors as shown in following Table 1.

The long-version personality questionnaire developed by IPIP-NEO (International Personality Item Pool Representation of the NEO PI-R) [6] containing 120 items is included in this study as the personality trait testing tool due to the following reasons:

1. NEO-PI-R's scales have proven to be useful tools in a number of applied fields.
2. IPIP representation is freely available in the public domain.
3. The instrument is relatively short (most people complete the questionnaire in 15–25 minutes) such that it is suitable for online testing and estimation.

2.3 Support vector machine

The Support Vector Machine (SVM) was derived from the Vapnik's structural risk minimization principle [5, 37] which is a novel machine learning algorithm for data classification and regression [11, 22]. It allows large information (training set) as a linear or nonlinear combination to divide data for classification [7]. The basic idea of SVM is to seek the optimal separating hyper-plane (support vectors) which is made up of elements with the ability to distinguish the data from the training set. Intuitively, the larger margin between two hyper-planes can get the higher classification accuracy [7]. The optimal separating hyper-plane has the maximal margin to each data group for separating the training data points into classes. It could effectively reduce the empirical risk based on the bounds of the generalization error, i.e. classification error on unseen examples [37].

SVM is one of useful supervised machine learning mechanisms based on classifier training, parameter validating, and performance testing [11]. The classifier training is to transform the training data into a higher dimensional feature space and the parameter validating is to find the optimal hyper-plane that maximizes the margin between different classes by different parameter settings such that an available classifier could be achieved. SVM could classify unknown data into the most appropriate data category/label according to the hyper-plane. Performance testing is based on the classification accuracy rate determination. One of the main attractions of using SVM is that it is capable of learning in sparse and high-dimensional feature spaces with very few training examples. Based on statistical learning algorithm, SVM has already been widely employed in many different research domains and real-world applications such as biological information analysis [26], intrusion detection of information security [1, 22, 33], image classification [38], and document categorization [15, 32]. So far, SVM has already become one of the standard tools for machine learning and data mining.

The simplest model of SVM is called the maximal margin classifier. As shown in (1), SVM attempts to place a linear boundary between the two different classes and to orient this line in such a way that the margin $1/\|w\|$ is maximized [7, 37]. SVMs derive a class decision by determining the separate boundary with maximum distance to the closest points, namely support vectors (SVs), of the training data set.

$$\begin{aligned} &y_i(w^T x_i + b) \geq 1, \\ &\quad \text{for } i = 1, 2, \ldots, N \text{ and } x_i \in A \cup B \qquad (1) \\ &\text{Minimize}_{w,b} \frac{1}{2}\|w\|^2. \end{aligned}$$

When two classes cannot be completely separated, this approach may not be feasible due to overlapping distribution. Therefore, the slack variable ξ is introduced to modify the maximum margin classifier that allows for misclassified examples. As shown in (2), the generalized optimal margin classifier, called soft margin classifier, can soften the hard constraint of separating cases completely. The soft margin classifier will choose a most possible class for the misclassified examples, also keeping the maximum margin to the nearest correct classification examples, if there is no hyper-plane that can clearly split examples. For more information

Table 1 Testing scores with the classification of five factors

Extraversion	Agreeableness	Neuroticism
Friendliness	Trust	Anxiety
Gregariousness	Morality	Anger
Assertiveness	Altruism	Depression
Activity Level	Cooperation	Self-Consciousness
Excitement-Seeking	Modesty	Immoderation
Cheerfulness	Sympathy	Vulnerability
Conscientiousness	**Openness**	
Self-Efficacy	Imagination	
Orderliness	Artistic Interests	
Dutifulness	Emotionality	
Achievement-Striving	Adventurousness	
Self-Discipline	Intellect	
Cautiousness	Liberalism	

about SVM, we recommend referring to [5, 7, 11, 34, 37].

$$\begin{aligned} & y_i(w^T x_i + b) \geq 1 - \xi, \\ & \quad \text{for } i = 1, 2, \ldots, N \text{ and } x_i \in A \cup B \\ & \text{Minimize}_{w,b} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}\xi_i, \\ & \quad \xi_i > 0, i = 0, 1, 2, \ldots, N. \end{aligned} \tag{2}$$

2.4 Issues of selecting parameters and kernel function in SVM

To apply SVM, the issues lie with the section of parameters and kernel function. These issues significantly influence the robust performance of SVM in many classification fields [1, 15, 26, 38]. Accordingly, the setting of kernel function and parameters both play an important role in SVM. As a result, how to choose a suitable kernel function and hyper-parameters for SVM becomes an important issue. In order to solve this problem, a few of solutions have been proposed. One of the common ways is to utilize cross-validation approach [7] in which a number of pairs of parameters are tested and the pair with which the highest accuracy would be picked.

As Chang and Lin [8] suggested, before building a classifier, determination of the kernel function and parameter pair must be done. In the kernel function selecting problem, there are four kernel functions in SVM, including linear, polynomial, radial basis function (RBF) and sigmoid functions. According to Smola's research, the RBF function is generally a reasonable first choice [34]. RBF kernel function uses the non-linear method to map samples to a high-dimensional space. In contrast to the linear kernel function, it can handle the relations among class labels and attributes that are non-linear [34]. Thus, the RBF was used as our kernel function of the SVM classification model.

Regarding the issue of hyper-parameter selection, there are two hyper-parameters used in the RBF kernel: kernel parameter γ and cost c. They are related to the complexity of SVM classifier and used to control the balance between maximizing the margin of separating hyper-plane and minimizing the classification error. The goal of parameter selection is to identify the best hyper-parameter pair (c, γ) so that the classifier can minimize the generalization error, generate accurate predictions, and mitigate the over-fitting problem [8, 27].

The problem of over-fitting in supervised machine learning is a situation that in order to minimize the generalization error of classifier [27] (i.e. the error made on the data used to train the classifier), the training algorithm is over trained to be a too precise model which leads to a very highly biased classifier [35]. This problem can considerably decrease the generalization accuracy when some land classes are not properly represented in the training data sets. To prevent over-fitting problem, conducting appropriate model complexity analysis is required. A too simple model could not learn the specificities of data but a too complex model will over learn the specificities of data, including outliers. The goal of cross-validation is to utilize all the available data for determining the complexity of model in order to get the balance between empirical error and generalization power (bias) of the classifier [35]. For this purpose, generally, cross-validation randomly divides the whole data into a number of folds. For each fold, the whole data sets except oneself are used as the training sets and the excluded one is used as a validating set. The mean error on each fold could be seen as the biased estimator. This is a good solution to select the parameters for getting the appropriate complexity of classifiers.

2.5 Multi-criteria decision-making

A Multi-Criteria Decision-Making (MCDM) problem is to find the best, compromised or optimal solution from all feasible alternatives evaluated on multiple and usually conflicting criteria, both quantitative and qualitative [21, 24, 25]. MCDM simplify the complex human decision-making process into the distance measure between criteria and alternatives. For multiple attribute decision making, TOPSIS can be an appropriate tool [19]. MCDM is employed to select a solution from several alternatives according to various criteria. To choose the qualified applicants in terms of several manager-defined capability preferences and personality factors is a MCDM problem. Therefore, the MCDM method can be applied to the recruitment process with complex and unintelligible information to support the decision of managers.

The technique of order preference by similarity to an ideal solution (TOPSIS) is a method of multi-criteria analysis model [20]. The basic idea of the TOPSIS is that each alternative may be viewed as an n-dimensional pattern. It is developed from the concept that the chosen alternative should have the shortest distance from the positive ideal solution (PIS) and the farthest from the negative-ideal solution (NIS) for solving a multiple-criteria decision-making problem [9]. In short, the ideal solution is composed of all best values attainable by the criteria, whereas the negative ideal solution is made up of all worst values attainable by the criteria. In addition, the preference among alternatives is ranked according to their similarity to the ideal solutions. The similarity is estimated by the Euclidean distance. In principle, the best alternative should be farthest from the NIS and closest to the PIS. However, an alternative can handle the shortest distance from the PIS, but not the farthest Euclidean distance from the NIS [39]. This article does not elaborate on the approach that has been successfully adopted in several studies. For the detailed processes of the TOPSIS method, refer to [20, 39].

Suppose that a MCDM problem has m alternatives denoted as $A_1, A_2, \dots, A_m$ from which decision-makers have to choose, and also n decision criteria denoted as $C_1, C_2, \dots, C_n$ with which alternative performance is measured, a typical MCDM problem can be expressed in matrix format D as below:

$$D = [x_{ij}]_{m\times n} = \begin{array}{c} \\ A_1 \\ A_2 \\ \vdots \\ A_m \end{array} \begin{array}{c} \begin{array}{cccc} C_1 & C_2 & \cdots & C_n \end{array} \\ \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix}_{m\times n} \end{array},$$
$$i = 1, 2, \dots, m;\ j = 1, 2, \dots, n, \tag{3}$$

where $x_{ij}, \forall i, j$ are the ratings of alternative A_i with respect to criterion C_j.

In order to transform the various criteria scales into an objective comparable scale, a normalized decision matrix R is obtained through linear scale transformation:

$$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^{m} x_{ij}^2}}, \quad i = 1, 2, \dots, m;\ j = 1, 2, \dots, n$$

$$R = [r_{ij}]_{m\times n} = \begin{array}{c} \\ A_1 \\ A_2 \\ \vdots \\ A_m \end{array} \begin{array}{c} \begin{array}{cccc} C_1 & C_2 & \cdots & C_n \end{array} \\ \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mn} \end{bmatrix}_{m\times n} \end{array}, \tag{4}$$
$$i = 1, 2, \dots, m;\ j = 1, 2, \dots, n.$$

The normalized method above preserves the property that the ranges of the normalized to [0,1].

Considering the different importance of each criterion, we can construct the weighted normalized decision matrix W as below:

$$v_{ij} = r_{ij} \times w_j, \quad i = 1, 2, \dots, m;\ j = 1, 2, \dots, n,$$

$$W = [v_{ij}]_{m\times n} = \begin{array}{c} \\ A_1 \\ A_2 \\ \vdots \\ A_m \end{array} \begin{array}{c} \begin{array}{cccc} C_1 & C_2 & \cdots & C_n \end{array} \\ \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ \vdots & \vdots & & \vdots \\ v_{m1} & v_{m2} & \cdots & v_{mn} \end{bmatrix}_{m\times n} \end{array}, \tag{5}$$
$$i = 1, 2, \dots, m;\ j = 1, 2, \dots, n,$$

where w_j denotes the relative importance with respect to x_j, and $w_1, w_2, \dots, w_n$ should satisfy $\sum_{j=1}^{n} w_j = 1$.

Let A^+ and A^- denote the PIS and the NIS, respectively. The distance of each alternative from A^+ and A^- can be currently calculated as below:

$$\begin{aligned} A^+ &= \left\{ \left(\max_{i=1,\dots,m} v_{ij} \mid j \in C_p \right), \left(\min_{i=1,\dots,m} v_{ij} \mid j \in C_n \right) \right\}, \\ A^- &= \left\{ \left(\max_{i=1,\dots,m} v_{ij} \mid j \in C_p \right), \left(\min_{i=1,\dots,m} v_{ij} \mid j \in C_n \right) \right\} \end{aligned} \tag{6}$$

C_p and C_n are the set of positive criteria (such as profit) and the set of negative criteria (such as cost) respectively. S_i^+ denotes the Euclidean Distance between the alternative and the PIS and S_i^- denotes the Euclidean Distance between the alternative and NIS.

$$S_i^+ = \sqrt{\sum_{j=1}^{n} (v_{ij} - v_j^+)^2}, \qquad S_i^- = \sqrt{\sum_{j=1}^{n} (v_{ij} - v_j^-)^2}, \tag{7}$$

for $i = 1, 2, \dots, m;\ j = 1, 2, \dots, n$.

In order to rank the order of all alternatives and the best solution can therefore be chosen from among a set of feasible alternatives, the positive performance index (PI^+) and negative performance index (PI^-) should be estimated as below:

$$PI_i^- = \frac{S_i^-}{S_i^+ + S_i^-}, \qquad PI_i^+ = \frac{S_i^+}{S_i^+ + S_i^-}. \tag{8}$$

3 The system framework

The proposed model in this article includes the following major steps as shown in Fig. 1. First, according to the different departments, we collect the personality trait data of existing employees and job applicants. In the proposed system, the training dataset denotes the original department human resource structure, which consists of existing employees. The testing data are utilized to indicate the new applicant or the existing employees who need internal job adjustment for the job vacancy. Second, the training datasets are used by the SVM to build the personnel prediction classification model of each department respectively. Third, the testing dataset was treated as the future unknown data and employed respectively to estimate the performance of every personnel model of departments. In this experiment, the business human resource model is built by existing employees to represent the personnel composition structure of the enterprise. We use the training dataset for training the SVM classification model to predict whether the job applicants are suitable for the position they applied. Fourth, according to the forecast results of SVM, we then add the factors which are the internal skilled assessment of enterprise as the criteria to process with the MCDM-TOPSIS method. Finally, according to the forecast results of SVM and TOPSIS, a recruitment suggestion is given to assist the managers in choosing which job applicant is most suitable and bring all positive factors into full play at work. Relying on the ability of TOPSIS, our proposed system will be able to estimate and rank the possible performance level of every applicant. The following figure describes the relevant research procedures.

3.1 Data collection and separation

In this step, every existing personnel and all job applicants are requested to fill out the personality questionnaire developed by IPIP-NEO. Since the SVM is a supervised machine learning method, firms have to collect the personality characteristic of existing personnel for training the department human resource model. Then, we separate these personality trait assessment reports by departments. That is, the personnel in each department model are separated according to their department. This can highlight the needs of personality traits of different departments and enable the system to predict accurately personnel personality trait.

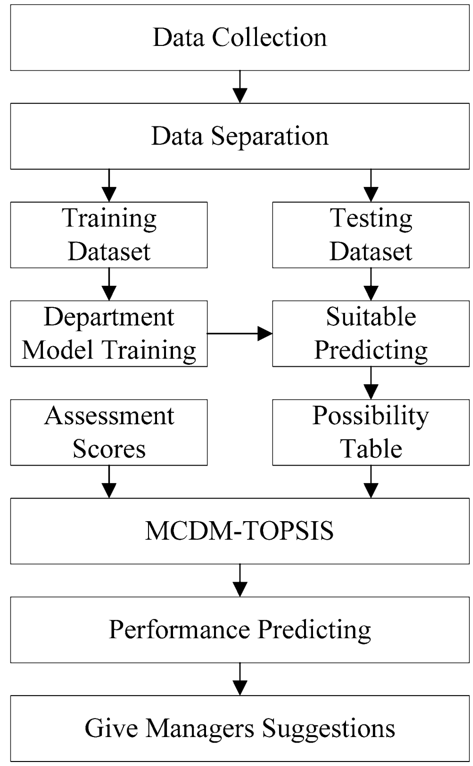

Fig. 1 Major steps of the model

3.2 Department model training

The model training procedure in this research is outlined in Fig. 2. First, in the data preprocessing step, the non-numerical data must be transformed into numerical data before the data were fed to SVM to process [8]. In order to conduct the two-class SVM experiment, we convert the non-numerical data of the department label to 1 if the employee belongs to the department which builds the specific model; otherwise, we convert the data to 0. Second, in order to increase the prediction accuracy rate and decrease the difference between the test data [7], data is normalized into interval [0, 1] in the normalization step before we advance to the next step. Third, we choose the RBF kernel function and cross-validation method to select the best parameters in this experiment. The rationale for choosing the RBF kernel function will be discussed in Sect. 4.3. The goal of the cross-validation method is to identify the best parameter pair (c and γ) so that the classifier can get the best prediction efficiency for unknown data samples. In this paper, the 10-fold cross-validation method is employed to determine the best pair of c and γ in the training dataset, thus yielding the best result for each respective department. Subsequently, this set of parameters is applied to validate the testing dataset. Finally, the training dataset is fed into the SVM for training each department model respectively.

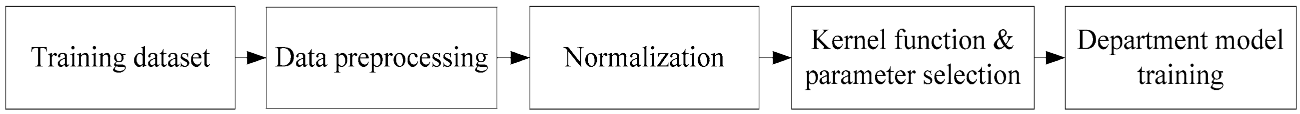

Fig. 2 Procedure of model training

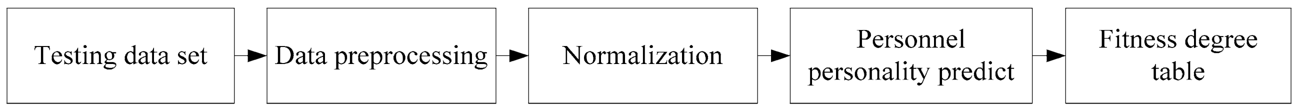

Fig. 3 Procedures of model testing

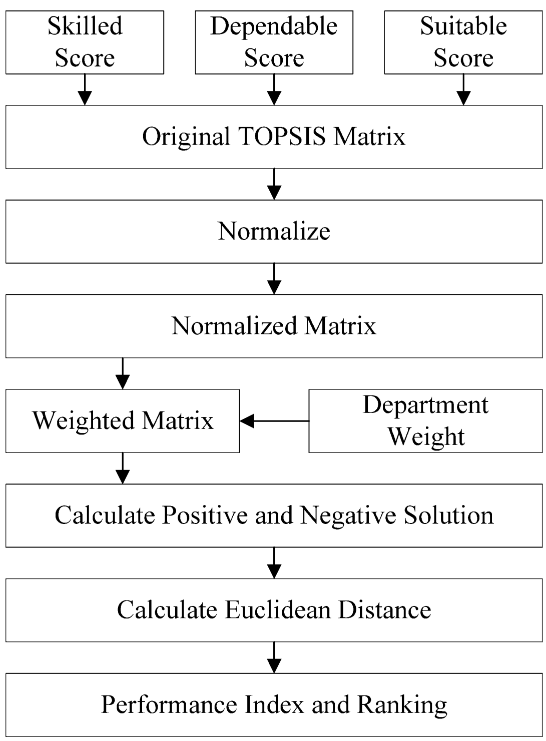

Fig. 4 Procedures of TOPSIS

3.3 Fitness prediction

In this section, the procedures of predicting personnel fitness are illustrated in Fig. 3. During the personnel prediction process, every personnel data must not only translate the non-numerical data into numerical data, but also normalize them into interval [0, 1]. After preprocessing and normalization, the dataset will individually be fed into each built department human resource model, which will produce two probabilities to indicate the degree of how the personnel belong to every data category, such as good and poor work performance. Our experiment focuses on the probability of good work performance and treats this probability as the department fitness degree for each personnel. That is, these department recruitment models are simulated as department managers and predict the fitness degree of each candidate for generating the recruitment suggestions. Finally, the proposed system integrates all the fitness degrees for managers as a basis for supporting personnel selection decision-making.

3.4 MCDM-TOPSIS

The TOPSIS procedures used in the proposed model are illustrated in Fig. 4. Detailed descriptions of each step are listed below.

Step 1: Original TOPSIS Matrix

The original matrix of TOPSIS consists of seven applicants and three criteria. That is, each applicant has the same criteria. By examining the difference in scores level of criteria, TOPSIS could be measured to discriminate between good and bad performance.

Step 2: Normalization and Normalized Matrix

In order to make the assessment criteria and form a basis for comparison, the original matrix has to be normalized.

Step 3: Weight and Weighted Matrix

In the measurement process, different criteria have different importance. After adding the weight, the TOPSIS will get a new weighted matrix. Before computing the ideal solution, the weighted value is computed.

Step 4: Calculate Positive and Negative Solution and Euclidean Distance

Calculate respectively the PIS and NIS in this step. After obtaining the PIS and NIS, calculate the Euclidean Distance for each alternative.

Step 5: Performance Index and Ranking

According to the result in Step 4, the positive performance index and negative performance index could be estimated. The best solution could be chosen accordingly.

Again, refer to the work of [20] and [39] for additional details of MCDM-TOPSIS implementation.

4 Experiment and results

4.1 Data description

The dataset used in this study is collected from a real business environment. Empirical dataset regarding the job performance of employees is given by an enterprise, which has already become the most scaled IC and LCD precision equipment manufacturer since 1978. The enterprise from which our used data collected is located in northern Taiwan and has many different service sites in central and southern Taiwan and China. In order to conduct this study, this firm provided their work performance scores of job assessment and the skill evaluation values of employees in seven different departments.

Our data collecting process is separated by the seniority of personnel. The training dataset consists of the senior personnel whose seniority is greater than one, while the testing dataset consists of the junior personnel whose seniority is less than or equal to one. The threshold of seniority was shown in Table 2.

The target firm was made up of 310 employees and 7 departments and there are totally 232 senior staffs, who have worked more than one year in the firm, included in all 7 department training data set. That is, totally 75% senior employees are engaged in the experiment to represent the existing human resource structure. The turnover rate per year of this firm approximates to 6%, that is, this firm approximately recruits 19 employees every year. In order to represent the real recruitment scenario, there are 14 (approximate to 74%) junior employees whose seniority in the firm is smaller than one year collected in a testing data set. The names of the seven departments and the total number of collected training data and testing data are respectively listed in following Tables 3 and 4.

Table 2 Data separation rule

Separation	Threshold
Training dataset	Seniority > 1
Testing dataset	Seniority <= 1

Furthermore, we invited the staffs to fill out a personality trait questionnaire test, the 120 questions edition of the IPIP-NEO. The dataset used in this experiment is made up of all the personality related factors which are collected and integrated form the questionnaire reports. Each personnel record in the dataset consists of the grades of 30 personality factors identified and determined by IPIP-NEO. The 30 personality factors can be regarded as the basic characteristics of a person. This dataset could be portrayed as the enterprise human resource structure and could be used to respectively construct the department models.

4.2 Data separation

According to the threshold established by the manager, we set the performance grade threshold as 80 degree. That is, if the original job assessment grade of an employee is greater than or equal to 80 degree, we identified him/her as good work performance. On the contrary, poor work performance would be identified, if the original assessment grade of an

Table 4 Test personnel detail description

Personnel	Seniority	Current department	Assess grade	Work performance
A	1	RD	75	Poor
B	1	FN	75	Poor
C	1	RD	80	Good
D	0.8	HR	78	Poor
E	1	FN	77	Poor
F	1	FN	68	Poor
G	0.5	IT	85	Good
H	0.8	ADM	88	Good
I	1	MC	80	Good
J	0.3	HR	65	Poor
K	1	HR	87	Good
L	0.8	FN	80	Good
M	0.7	MC	78	Poor
N	1	RD	78	Poor

Table 3 Name and number of data of seven departments

Department name	Number of senior employee data	Number of junior employee data
Administration (ADM)	32	1
Material Control (MC)	35	2
Project Management (PM)	28	0
Research & Development (RD)	38	3
Information Technology (IT)	37	1
Human Resource (HR)	29	3
Finance (FN)	33	4

employee is lower than 80 degree. The threshold of work performance and the number of employees who got the specific performance in the training set were shown in Table 5. While training the department recruitment models, we labeled good work performance as "0" and labeled poor work performance as "1" for the collected data to learn the relationship between personality factors and work performance. Notice that the same data separation rule is also applied to the testing data set as shown in Table 4.

4.3 Kernel and parameter selection

In machine learning algorithms, the perfect separation may not be possible, or it may result in a model with so many feature vector dimensions that the model does not generalize well to other data. The problem is known as over-fitting. The m-fold cross-validation is one of the most effective methods to avoid over-fitting based on the findings of [4, 27]. Besides, the identified parameters could be applicable to a specific test dataset which is quite objective, and could help to prevent the over-fitting problem. For these reasons, the m-fold cross-validation parameter selection method is used to search the parameters in our experiment. We set both c and γ as two numerical sets in our parameter validation process. In the m-fold cross-validation, the training data is divided randomly into m folds (sub-datasets) of equal size. The $m-1$ sub-datasets are employed to train the classifier. Sequentially, the remaining 1 fold is used for validating the accuracy of the classifier. There are m experiments conducted in total for validating a pair of parameters. The cross-validation accuracy is formatted as the percentage of correctly classified data. Finally, the effectiveness of these parameter pairs can be determined by the average accuracy.

In order to obtain a better performance, the impact of different kernel functions and parameter settings are compared. Depending on the result of [27], the generalization error of 10-fold cross-validation is smaller than those of others. In this experiment, the 10-fold cross-validation method is used to identify the best parameters for different kernel functions, such as linear, polynomial, RBF, and sigmoid kernels, for our SVM estimator. Different kernel functions have different parameters need to be selected as listed in the Table 6. The parameter *degree* in the polynomial kernel function belongs to the set $\{1, 2, 3\}$ for validating the highest accuracy. In our experiments, the accuracy rate changes slightly in the polynomial kernel when degree is larger than 3. And, the

Table 5 The threshold of work performance

Department	Number of collected data	
	Good work performance (Assess grade >= 80)	Poor work performance (Assess grade < 80)
ADM	20	12
MC	21	14
PM	15	13
RD	26	12
IT	27	10
HR	19	10
FN	24	9

Table 6 The best hyper-parameter of each department

Department model	Linear kernel Accuracy	c	Polynomial kernel Accuracy	c	γ	*Degree*
ADM	80.6452	1000	90.3226	1	0.01	3
FN	87.0968	2	90.3226	100	2	1
HR	83.871	50	93.5484	750	2	1
IT	87.0968	2	93.5484	50	2	2
MC	80.6452	1000	80.6452	10	1	2
PM	80.6452	10	83.871	750	0.001	3
RD	83.871	1	87.0968	250	0.001	1

Department model	RBF Kernel *Accuracy*	c	γ	Sigmoid Kernel *Accuracy*	c	γ
ADM	**92.3077**	750	0.5	90.3226	100	0.001
FN	**92.3077**	2	0.5	90.3226	750	0.1
HR	**96.1538**	10	0.01	90.3226	50	0.01
IT	**96.1538**	750	0.0001	93.5484	500	1
MC	**84.6154**	50	0.1	80.6452	50	0.0001
PM	**84.6154**	250	0.1	80.6452	250	0.0001
RD	**88.4615**	500	0.01	80.6452	750	0.1

parameter *coef* in polynomial and sigmoid kernels does not significantly affect results. The default value setting of this parameter is zero in both kernels.

The parameters tried in the validation process are $\gamma \in \{2, 1, 0.5, 0.1, 0.01, 0.001, 0.0001\}$ and $c \in \{1000, 750, 500, 250, 100, 50, 10, 2, 1\}$. Then, we perform a cross-validation process for each parameter pair (γ, c) to obtain the highest classification accuracy. That is, we first set $\gamma = 2$ and $c \in \{1000, 750, 500, 250, 100, 50, 10, 2, 1\}$ to validate all pairs such as $\{(2, 1000), (2, 750), \ldots, (2, 1)\}$, then sequentially set $\gamma = 1$ until all pairs are validated. Finally, the parameter pair with the highest accuracy would be picked. The highest performance of different department recruitment model with various kernel functions and parameter settings are displayed in Table 6.

Since the polynomial and RBF kernels could handle the nonlinear relation between class labels and attributes [8, 34], how the performances of each department classifier built by polynomial and RBF kernels outperform others' can be observed. Although the polynomial kernel also could handle the nonlinear relationship between labels and attributes, it has more hyper-parameters (4 parameters: c, γ, *degree* and *coef*) than the RBF kernel (2 parameters: c and γ) also, the model training time is much longer than RBF kernel, and the accuracy is lower than RBF kernel.

Theoretically, the accuracy of an SVM model is largely dependent on the selection of the model parameters such as c and γ. Dependent on data distribution, these two parameters control the tradeoff between allowing training errors (separating accuracy) and forcing the margins of hyper-planes that permit some misclassifications. In general, a more complex classifier will adopt larger parameter values. Practically, the skills and personality traits of employees from different departments should be different. Varied weights for different personality and skill factors are considered in hiring, reallocating, and performance assessment. Due to the staff personality structure variation, the complexity of each department model would be different. In our experiment, different departments also consider different personality types/traits in the recruitment process. For instance, department ADM (Administration) particularly considers the factors belonging to five personality factor categories ("Extraversion", "Agreeableness", "Neuroticism", "Conscientiousness", and "Openness"). However, the main task of department FN (Finance) is related to company accounts. The manager especially focuses on the factors belonging to two personality factor categories ("Conscientiousness" and "Agreeableness"). The recruitment decision of ADM department will be more complex, compared with the recruitment decision of FN (Finance). We can also observe that the values of c and r in the ADM department are higher than those in FN department.

4.4 Results of SVM personality prediction

We utilize these built department recruitment models for the managers to obtain the fitness degree of each candidate for supporting the recruitment decision. The process to achieve the fitness degree of each candidate in each department is described as follows. After an individual junior employee data is fed into each built department recruitment model, the SVM model will output the probability (fitness degree) that the applicant would get a good work performance in each department. The higher probability of getting a good work performance in a department intuitively represents that the applicant has a higher fitness level in working in a department.

The department fitness degree of applicants would be individually judged by each department model. Table 7 shows the fitness degree between the 14 junior employees and various departments. Each numerical value in the table represents the fitness degree of the personnel to work in a specific department. For example, employee A has 7.7% fitness for working in the ADM department and 35.5% fitness in the IT department.

In accordance with the department fitness degree, this model provides the fitness table to help managers with making human resource decisions. For instance, which job-hunter would be most suitable to hire for which department or the employee who got lower work performance should be reassigned to which position.

Let us see the following two cases:

Case 1: Assume the firm is about to hire new employees of IT department, and there are 14 applicants (identified as $A \sim N$) applying for this job. If this firm plans to hire two employees, A and G will be nominated according to the table of fitness degrees. Note that since G is a new incoming employee, his/her work performance has not been evaluated by the firm while we collect the questionnaire report. The assessment grade will be estimated later when his/her seniority is enough. The estimating result is totally corresponding to our prediction.

Case 2: Assume the firm plans to transfer F in the department FN to another department because his/her assessment grade is merely 68. According to the table of fitness degree, the proposed model will suggest the manager to transfer F from FN to MC department.

So far, based on personality questionnaire and SVM, our proposed system can automatically score the fitness between job's personality requirement and candidate's personality trait in addition to the conventional Job-CV matching, and then generates the best set of job candidates.

4.5 Results of TOPSIS performance prediction

Extend the cases above; assume that the managers actually transferred the employees to the specific departments. Based

Table 7 Fitness degree

	ADM	FN	HR	IT	MC	PM	RD
A	0.077	0.087	0.065	**0.355**	0.214	0.247	0.094
B	0.098	0.067	0.067	0.064	0.165	0.143	**0.389**
C	0.105	0.111	0.079	0.022	0.153	0.115	**0.643**
D	0.107	0.085	0.085	0.042	0.188	0.161	**0.239**
E	0.108	0.088	0.062	0.029	**0.421**	0.137	0.111
F	0.111	0.001	0.038	0.094	**0.147**	0.124	0.011
G	0.082	0.092	0.057	**0.699**	0.139	0.113	0.160
H	0.056	**0.78**	0.078	0.046	0.136	0.094	0.049
I	0.054	0.065	0.15	0.044	**0.362**	0.094	0.075
J	0.141	0.091	0.012	0.143	**0.245**	0.232	0.005
K	0.054	0.062	**0.76**	0.076	0.136	0.094	0.116
L	0.142	**0.214**	0.005	0.177	0.091	0.165	0.118
M	**0.197**	0.096	0.071	0.094	0.141	0.163	0.041
N	0.14	**0.226**	0.055	0.139	0.091	0.088	0.185

Table 8 Performance ranking index table

	ADM	FN	HR	IT	MC	PM	RD
A	0.202	0.131	0.12	**0.502**	0.348	0.365	0.171
B	0.197	0.097	0.099	0.103	0.207	0.345	**0.597**
C	0.186	0.146	0.111	0.077	0.101	0.17	**0.942**
D	0.193	0.113	0.116	0.075	0.203	0.228	**0.363**
E	0.195	0.117	0.085	0.059	**0.853**	0.057	0.163
F	0.168	0	0.043	**0.105**	0.027	0.046	0
G	0.212	0.14	0.117	**1**	0.255	0.045	0.259
H	0.233	**0.922**	0.096	0.107	0.132	0.157	0.099
I	0.123	0.081	0.191	0.038	**0.837**	0.044	0.131
J	0.179	0.115	0.009	0.199	0.483	**0.521**	0
K	0	0.078	**0.905**	0.097	0.132	0.168	0.198
L	0.166	**0.404**	0	0.238	0	0.02	0.201
M	**0.329**	0.121	0.119	0.117	0.251	0.088	0.092
N	0.208	**0.419**	0.067	0.19	0.084	0.038	0.301

on the results of SVM personality fitness prediction, we then utilize the standard TOPSIS procedure to predict whether the employees can get better performance in the suggested departments than in the original departments. In this proposed model, there are three criteria included in the matrix such as skilled score, dependable score, and fitness score. The skilled score and dependable score are measured by internal assessment of enterprise. The fitness score is obtained from the result of SVM personality predicting. Based on these three criteria, the performance ranking index of each department will be given by MCDM method, respectively. Each department's performance ranking indexes listed as following Table 8.

Comparing with Table 7, it is easy to see the results are basically matched. A comparison of these two tables, Tables 7 and 8, shows that an employee could get possible higher work performance in the higher fitness department suggested by SVM. However, only the personality and working performance predicting results are not matched on employee F and J. After further re-checking, we found the results of the initial personality questionnaire show that these two employees did not totally finish it. The designed result of each questionnaire should be equipped with five dimensional scores respectively, but only three are found in the employee's questionnaire and the remaining two dimensional scores are all zero. This situation means the result cannot be displayed correctly. In addition, our proposed double-layered employee choosing system provides an error-checking function for helping firms avoid assigning inappropriate jobs to employees.

4.6 Evaluations and discussions

In our experiment, employee *G* is totally new employee of the IT department. He/she has not yet been involved in any internal performance measurement from the firm in the data collection period. Fortunately, after the performance prediction, employee *G* has already had his first assess grade 85 showed in Table 4. Comparing with our performance ranking indexes Table 8 (employee *G* is the one who is most suitable work in IT department) and the firm identified work performance threshold (assess grade $\geq$ 80 indicates good work performance), our proposed predict mechanism actually got accurate results.

After the experiment, we also interviewed with executives of the human resource department of the firm. According to the experimental results, the executives thought, in comparison to their existing personnel work performance measurement, our proposed mechanism performs better. Our proposed model can avoid employing inappropriate employees due to people's selfish motives. The model makes the recruitment more fair and objective. Based on the experimental results and personnel manager's opinions, we believe that the proposed human resource prediction framework is really beneficial for the business. In particular, better quality of recruitment could significantly reduce the employees training cost (e.g. time and expense) and enable employees to be skilled at their work position faster. In addition, we found that the fitness between the characteristics of job and the personality trait of candidate is more important than the skilled score of candidate when an enterprise would like to recruit a new employee.

Comparing the results of SVM fitness degree prediction in Table 7 with the results of TOPSIS prediction that joins skilled score in Table 8, we found that only personnel *F* got the different prediction. This shows the personality is a key factor for impacting whether a person is suitable for a particular job. This is different from traditional thinking that the skilled score is more important than personality. A partial explanation for this may lie in the fact that personnel could get the necessary skill of job by general training but the necessary personality is hard to be trained.

Although our findings show that the work performance is closely related to the personality of an employee, the effects of personality are cumulative and compound over time. Because they have to face the pressure, knowledge, skills, and abilities of their work, the personality of personnel might be changeable. Therefore, firstly, to continuously collect the personality trait data of different departments in this framework would be needed, in order to distinguish the personality type of work performance significantly, including high and low work performance. Secondly, in order to make more accurate predictions, the job applicants and the employees who get low work performance and need to relocate work position should answer the questionnaire while they apply for the work or relocate to another position.

While this body of research has the demonstrable merit of offering valuable insights into personnel recruitment based on 232 available and usable personality test reports included in the training dataset, the enterprise can keep collecting more personality and performance data of personnel to improve the prediction accuracy. Particularly, that would be able to efficiently improve the accuracy of SVM classifier for the fitness between personality and job position predicting.

5 Conclusion

Enterprises are now in an era of global talent search. The primary motivation behind this study was in the development of a qualitative recruitment system that could help enterprises recruit the right personnel for the right position. Because of limited resources on recruiting budgets, accurate candidate employment is indeed a managerial issue. Therefore, it is vital to develop a useable and reliable model for predicting and discovering talented employees. The present study enhances the previous studies' findings by providing a much more detailed examination of personnel recruitment.

In this study, the model is a combination of the credible questionnaire developed by IPIP-NEO, SVM classifier, and TOPSIS. Unlike existing methods, we use the SVM with probability estimates to first obtain the most likely classes that represent the fitness degree of candidates and then determine the possible work performance based on the TOPSIS. First, the questionnaire was used to analyze the personality of personnel. The personality trait was utilized as job fitness measurement criteria of personnel or job applicant. Second, the fitness degree between the personnel's personality type and department is estimated by SVM. Finally, based on personality fitness, joining skill score and dependable score given by firms' internal assessment, TOPSIS method is used for determining and ranking the possible work performance level of each candidate. The position adjustment or recruitment suggestion would be given to managers. Most of the CEOs thought that to find, hire, and retain a qualified employee is the greatest challenge for operating enterprises and the most important factor contributing to company growth. We believe this proposed model can not only be used to predict suitable candidate for the vacancy but also the possible job performance estimation after the candidate actually takes up the vacancy. Furthermore, the proposed model could be used for job recruitment as well as for decanting the employees' work performance.

There are several directions for extending this research: (i) what is the relationship between the big five factors and the company's existing work performance measurement factors? Future work can hopefully clarify this important relationship concern. (ii) How to develop a more efficient way to determine the parameters of the SVM model when the human resource structure in a business division has been changed? Appropriate parameter selection is a critical task for improving the prediction effectiveness of the model.

Acknowledgements This research was supported by the National Science Council of Taiwan (Republic of China) under the Grant NSC 96-2416-H-009-017.

References

1. Ambwani T (2003) Multi class support vector machine implementation to intrusion detection. In: Proceedings of the international joint conference of neural networks, vol 3, pp 2300–2305
2. Barrick MR, Mount MK (2003) Yes, personality matters: moving on to more important matters. Human Perform 18(4):359–372
3. Beagrie S (2005) Personnel today—how to excel at psychometric assessments. Available at http://www.personneltoday.com/articles/2005/03/22/28745/psychometrics-how-to-excel-at-psychometric-assessments.html. Accessed 23 Dec 2007
4. Bengio Y, Grandvalet Y (2004) No unbiased estimator of the variance of k-fold cross-validation. J Mach Learn Res 5:1089–1105
5. Boser BE, Guyon IM, Vapnik VN (1992) A training algorithm for optimal margin classifiers. In: Proceedings of the fifth annual workshop on computational learning theory pp 144–152
6. Buchanan T, Johnson JA, Goldberg LR (2005) Implementing a five-factor personality inventory for use on the Internet. Eur J Pers Assess 21(2):116–128
7. Burges CJC (1998) A tutorial on support vector machine for pattern recognition. Data Min Knowl Discov 2(2):121–167
8. Chang CC, Lin CJ (2003) LIBSVM: A library for support vector machines, 2001. Software and documents available at http://www.csie.ntu.edu.tw/~cjlin/libsvm. Accessed 10 Feb 2008
9. Chen MF, Tzeng GH (2004) Combining grey relation and TOPSIS concepts for selecting an expatriate host country. Math Comput Model 40(13):1473–1490
10. Costa PT Jr, McCrae RR (1992) NEO PI-R professional manual. Psychological Assessment Resources, Odessa
11. Cristianini N, Shawe-Taylor J (2000) An introduction to support vector machines and other kernel-based learning methods. CUP, Cambridge
12. Deloitte, Touche (2006) Technology fast 500 CEO survey results. The Economist Intelligence Unit, The CEO's role in talent management, United States, 2006
13. Dunn WS, Mount MK, Barrick MR, Ones DS (1995) Relative importance of personality and general mental ability in managers' judgments of applicant qualifications. J Appl Psychol 80(4):500–509
14. Goodstein LD, Lanyon RI (1999) Applications of personality assessment to the workplace: a review. J Bus Psychol 13(3):291–322
15. He J, Tan AH, Tan CL (2003) On machine learning methods for Chinese document categorization. Appl Intell 18(3):311–322
16. Hong WC, Pai PF, Huang YY, Yang SL (2005) Application of support vector machines in predicting employee turnover based on job performance. In: Lecture notes in computer science, vol 3610. Springer, Berlin, pp 668–674
17. Hough LM (1998) Personality at work: issues and evidence. In: Hakel M (ed) Beyond multiple choice: evaluating alternative to traditional testing for selection. Erlbaum Assoc, Hillsdale
18. Hough LM, Oswald FL, Ployhart RE (2001) Determinants, detection, and amelioration of adverse impact in personnel selection procedures: Issues, evidence, and lessons learned. Int J Sel Assess 9(1):152–194
19. Hu YC (2008) Classification performance evaluation of single-layer perceptron with Choquet integral-based TOPSIS. Appl Intell 29(3):204–215
20. Hwang CL, Yoon K (1981) Multiple attribute decision making. Springer, Berlin
21. Kuo MS, Tzeng GH, Huang WC (2007) Group decision-making based on concepts of ideal and anti-ideal points in a fuzzy environment. Math Comput Model 45(3–4):324–339
22. Laskov P, Düssel P, Schäfer C, Rieck K (2005) Learning intrusion detection: supervised or unsupervised? In: Lecture Notes in Computer Science, vol 3617. Springer, Berlin, pp 50–57
23. Li S (2004) A behavioral choice model when computational ability matters. Appl Intell 20(2):147–163
24. Li DF (2007) Compromise ratio method for fuzzy multi-attribute group decision making. Appl Soft Comput 7(3):807–817
25. Li YM, Kao CP (2009) TREPPS: a trust-based recommender system for peer production services. Exp Syst Appl 36(2):3263–3277
26. Maglogiannis I, Zafiropoulos E, Anagnostopoulos I (2007) An intelligent system for automated breast cancer diagnosis and prognosis using SVM based classifiers. Appl Intell 30(1):24–36
27. Markatou M, Tian H, Biswas S, Hripcsak G (2005) Analysis of variance of cross-validation estimators of the generalization error. J Mach Learn Res 6:1127–1168
28. McHenry JJ, Hough LM, Toquam JL, Hanson MA, Ashworth S (1990) Project A validity results—the relationship between predictor and criterion domains. Pers Psychol 43(2):335–354
29. Mount MK, Barrick MR (1995) The Big Five personality dimensions: implications for research and practice in human resource management. Res Pers Human Resour Manag 13:153–200
30. Mount MK, Witt A, Barrick MR (2000) Incremental validity of empirically-keyed biographical scales over GMA and the big five personality constructs. Pers Psychol 53(2):299–323
31. Periatt JA, Chakrabarty S, Lemay SA (2007) Using personality traits to select customer-oriented logistics personnel. Transp J 46(1):22–37
32. Schmidt FL, Hunter JE (1998) The validity and utility of selection methods in personnel psychology: practical and theoretical implications of 85 years of research findings. Psychol Bull 124(2):262–274
33. Shon T, Moon J (2007) A hybrid machine learning approach to network anomaly detection. Inf Sci 177(18):3799–3821
34. Smola AJ (1998) Learning with kernels. PhD thesis, Department of Computer Science, Technical University Berlin, Germany
35. Sterlin P (2007) Overfitting prevention with cross-validation. Master thesis, University Pierre and Marie Curie (Paris VI): Paris, France
36. Tzeng HM, Hsieh JG, Lin YL (2004) Predicting with a support vector machine nurses' intention to quit: a new approach to set up an early warning mechanism in human resource management. Comput Inform Nurs 22(4):232–242
37. Vapnik VN (1995) Statistical learning theory. Wiley, New York
38. Wong WT, Shih FY, Liu J (2007) Shape-based image retrieval using support vector machines, Fourier descriptors and self-organizing maps. Inf Sci 177(8):1878-1891
39. Yoon KP, Hwang CL (1995) Multiple attribute decision making: an introduction. Sage, London

Yung-Ming Li is an Associate Professor at the Institute of Information Management, National Chiao Tung University in Taiwan. He received his Ph.D. in Information Systems from the University of Washington. His research interests include Peer-to-Peer networks, Internet economics, and business intelligence. His research has appeared in IEEE/ACM Transactions on Networking, Decision Support Systems, Electronic Commerce Research and Applications.

Cheng-Yang Lai is a Ph.D. student at the Institute of Information Management, National Chiao Tung University. He received his B.S. and M.S. degrees in Department of Information Management from the Chung Hua University, Taiwan. His research interests include recommender systems, social computing and electronic commerce. His research has appeared in Hawaii International Conference on System Sciences (HICSS).

Chien-Pang Kao received his M.S. degree from Institute of Information Management, National Chiao Tung University, Taiwan and B.S. degree in Management Information Systems from the National Yunlin University of Science and Technology, Taiwan. His research interests focus on recommender systems, business intelligence and electronic commerce. His research has appeared in Expert Systems with Applications.