International Journal of
Information
and
Management
Sciences

http://ijims.ms.tku.edu.tw/main.php

International Journal of
Information and Management Sciences

33 (2022), 333-352. DOI:10.6186/IJIMS.202212_33(4).0004

Prediction of Dynamic Job Performance in Human Resource Management Using a Hybrid Data Mining Scheme-A Case of the Fast Fashion Industry

Wen-Chin Chou, Chun-Chi Yang and Chi-Jie Lu[*]
Fu Jen Catholic University

Keywords

Human resource management
job performance prediction
data mining, variable selection
hybrid model

Abstract.

Human resource management (HRM) enables organizations to recognize the dynamic complexity of the environment. However, HRM practices require a multidisciplinary approach to allow feedback between organizational actions and the results, which increases competitive advantage. This study breaks away from the traditional static model development at a given time. Instead, we collect three years of data and use a hybrid data mining scheme to create a practical dynamic job performance prediction model. The knowledge extracted from the three years model shows that dynamic views exist as the organization faces changes, which the previous single static model cannot explain. We also find that although dynamic views exist in the models, the effects of several critical variables persist, and the importance varies with the year. Finally, we propose suggestions for improving HR practices from the extracted knowledge, thereby promoting the strategic value of HR management.

1. Introduction

Organizational performance is the result of an organization's resources. Dynamic capabilities are essential for organizations to yield superior performance in changing environments and are the key to competitive advantages (Teece et al. [51]). Past empirical studies have proven that human resource management (HRM) significantly impacts organizational performance (Becker and Gerhart [10], Delaney and Huselid [16]). HRM practices attract, motivate, and retain employees to ensure an organization's survival (Schuler and Jackson [47]). HRM practices contribute to organizational performance by developing different aspects of human resources to generate higher cost efficiencies, reduced turnover, and better and lower recruitment and training costs (Backhaus et al. [8],

[*]corresponding author

Peterson [41]). Therefore, human resource decisions help create and sustain organizational capabilities and competitive advantages (Becker and Gerhart [10]).

Performance management is one of the most critical HRM practices. Job performance reflects scalable actions, behavior, and outcomes that employees engage with or contribute to within organizations (Campbell [14]). Companies that seek competitive advantages through employees must be able to manage the behavior and performance of all employees (Noe et al. [38]). Performance management can verify HRM practices such as personnel selection, training, and job design, and performance results can incorporate into the feedback of merit pay systems. Therefore, performance management can effectively promote the HRM cycle, helping to create and maintain organizational capabilities and competitive advantages.

HRM has become a new field of data mining research (Strohmeier and Piazza [50]). Data mining in job performance prediction applies widely across industries. Such as information technology (Al-Radaideh and Al Nagi [2]), banking (Bach et al. [7], Honglei [26]), insurance (Delgado-Gómez et al. [17]), customer service center (Valle et al. [52]), and the textile industry (Wang and Shun [56]). However, by analyzing one hundred relevant studies, Strohmeier and Piazza [50] found that most studies focused on developing data mining algorithms and seldom discussed the implications and practices of relevant variables. These studies failed to integrate with human resource disciplines to benefit human resource practices. In addition, the primary mining data is cross-sectional at a given time. The relevant variables used are limited, and the data collected from multiple periods are rarely applied. Therefore, these studies do not reflect that human resource strategy adjusts dynamically because of environmental change, and organizational learning to respond is a dynamic process (Nevis et al. [36]). So, they cannot, in turn, generate practicable knowledge and decision-making information.

In the fast fashion industry, the market is changing rapidly, which requires participants to be more flexible and responsive (Christopher et al. [15]). The fast fashion industry has many unique characteristics (Goto and Endo [21]): shorter product life cycles, high product variability, subjective evaluation of product appearance, long and complex operations, high manual labor, and high labor cost. Most of the characteristics are related to people. Therefore, HRM is one of the core issues in the industry. Finding the critical variables that affect job performance and improving human resources performance is essential to competitive advantage in the long run.

This study collects information on the performance of new employees from a leading fast fashion company in Taiwan for three years. Moreover, this study proposes a hybrid data mining scheme by using two variable sections, namely the information gain method (IG) and ridge regression method (RR), and five data mining prediction techniques, namely stepwise regression (SR), M5P, random forest (RF), support vector regression (SVR), and backpropagation neural network (BPN). This study uses it to analyze individual-level performance-related data over a long period to fill the gaps of previous studies, thereby benefiting HRM practices.

This study constructs a job performance prediction and analysis framework to achieve two aims: (1) To develop dynamic performance prediction models by collecting and analyzing cross-time dimensions and diverse variable data to compensate for the deficiencies of previous static model studies, and (2) To develop prediction models for performance comparison using variable selection methods and different algorithms; to conduct an in-depth discussion of critical variables and extract knowledge as decision-making information for promoting HRM practices.

The rest of this paper organizes a brief review of the previous studies on individual-level variables related to job performance and data mining in performance prediction models conducted in Section 2. The proposed hybrid data mining scheme describes thoroughly in Section 3. Section 4 presents the empirical results. In Section 5, we explore the best model variables, extract critical information and knowledge, and provide practice suggestions. Finally, this study's conclusion and research limitations conduct in Section 6.

2. Literature Review

2.1. Job performance

Job performance is a central construct in industrial/organizational psychology (Campbell [14], Murphy and Cleveland [35]). Campbell [14] described job performance as employees' behavior when fulfilling organizational expectations and prescribed or official role requirements. Borman and Motowidlo [12] described job performance as all behavior related to organizational goals, and the behavior can measure according to the degree of individual contribution to the goals. According to these descriptions, job performance is when individual goals meet organizational goals. It is also the combination of individual and organizational performance to achieve goals and improve overall organizational effectiveness.

The purposes of performance management systems are of three kinds: strategic, administrative, and developmental (Noe et al. [38]). Strategic purpose implements through defining the characteristics, abilities, behavior, and results of employees required to carry out the strategies, developing measurement and feedback systems to connect employees' activities with organizational goals. The administrative purpose is to use performance results in salary management, promotion, layoffs, and other administrative decisions. The developmental purpose is the development of employee abilities or skills through performance management results so that employees can perform their tasks more effectively. Therefore, the management of job performance can not only improve human resource practices but also promote the achievement of organizational strategies.

2.2. Personal variables affecting job performance

Job performance is influenced by individual characteristics, outcomes, and work environment (Waldman and Spangler [55]). The relevant variables center on employee characteristics and critical behavior during the performance period. Such as level of education(Ng and Feldman [37]), work experience (Quiñones et al. [43]), age (Saks and Waldman [45], Waldman and Avolio [54]), intelligence(Ree et al. [44]), personality (Kristof-Brown et al. [31]), overtime(Schaufeli et al. [46]), and absenteeism(Judiesch and Lyness [27]). Furthermore, according to the cognitive dissonance theory (Festinger

[18], Festinger and Aronson [19]), when a person holds two psychologically inconsistent thoughts, he or she experiences discomfort and attempts to reduce the dissonance. The most common way to reduce dissonance is to make cognitions more consistent with each other (consonants). Therefore, employees will adjust their attitudes and behaviors to avoid cognitive dissonance when they recognize their work results as inconsistent with supervisor ratings. Therefore, this study incorporated performance, potential, and promotion records as rated by supervisors in the previous year into next year's performance predictors.

2.3. Data mining in job performance prediction

By taking the support vector machines approach, Delgado-Gómez et al. [17] developed performance prediction models based on intelligence and personality tests of insurance agents, reducing the direct and indirect costs of recruitment, training, and personnel maintenance. Aluko et al. [4] use logistic regression and support vector machine techniques to predict the academic performance of architecture students based on prior academic achievement information. The developed model can help universities as a decision-making tool in selecting students for admission. By taking the decision tree approach, Al-Radaideh and Al Nagi [2] collected information on the academic qualifications and appointment of employees in three information technology companies. They built performance prediction models, which can use for the performance prediction of recruits. Arfaee et al. [6] cooperated with the State Administration of Taxation to collect personal information, work information, and training information of tax administration assessors. They follow the CRISP data mining methodology and use decision tree and neural network technologies to develop prediction models to determine the key variables affecting job performance. The model is used to assess employees' current educational situation and as decision-making information for training strategies. By taking the multiple regression approach, Pejić Bach et al. [40] constructed performance prediction models based on the personality traits of bank personnel and applied them to recruitment. By taking the neural network approach, Wang and Shun [56] collected information on the work experience, family status, expertise, teamwork skills, and stress tolerance of employees in the textile industry. They built employee performance prediction models and applied them to employee retention, which can reduce the cost of retention.

According to the above studies, data mining in performance management is widespread across industries, and various methods exist. The data used include employee background information such as age, gender, family, education, personality, work experience, training, and tenure. However, these studies only incorporated some of the data, and few of them integrated a variety of information and conducted long-term data collection and analysis.

3. Proposed Job Performance Prediction Scheme

3.1. Variable selection methods

This study adopts two variable selection methods: The information gain method (IG) and the ridge regression method (RR). IG is a filter feature selection method to measure features to the dataset of the impact on entropy value. It is suitable for lower-dimensional datasets of variable selection (Liu et al. [33]). IG can fully consider the relationship between variables to evaluate the correlation between variables and predicted values. This measure ranks the importance of the variable under study via the information obtained from a dataset, removing the variable to fall below ranking or relatively low IG values. IG plays an important role in the field of variable selection. It is widely used to examine the relationships between variables and variables (Abelln and Castellano [1], Alam et al. [3], Liu et al. [33]).

RR has been widely applied in many studies and has excellent variable selection performance (Arefeen et al. [5], Barry et al. [9], Kim et al. [28]). RR and least absolute shrinkage and selection operator (Lasso) are both extensions of the linear regression method and have similar basic concepts, and both add penalty parameters to limit the model coefficients, minimizing the complexity of the model. RR and Lasso main difference is that Lasso integrates the least absolute selection and shrinkage operator with L1 regularization, which can force compression of the coefficients of covariates, making a minor contribution to the model to exactly zero to attain lower variance to reduce the problem of overfitting (Hastie et al. [22], Kwon et al. [32]). While RR uses the L2 regularization technique to shrink model coefficients in RR, which addition of appropriate L2 penalties to the model shrinks all the coefficients to a nonzero value or a value approaching zero and then minimizes the sum of squared error, and further controls the trade-off between bias and variance to reduce overfitting (Hoerl and Kennard [23]). However, the lasso penalty expects many coefficients to be zero. Thus, the Lasso is not robust to high correlations among predictors and will arbitrarily choose one and ignore the others (Friedman et al. [20]). RR performs well with many predictors, each having a small effect, and prevents coefficients of linear regression models with many correlated variables from being poorly determined and exhibiting high variance (Ogutu et al. [39]). RR can improve the prediction performance for multicollinearity problems. Thus, RR is an ideal method than Lasso on variable selection in this study.

3.2. Data mining methods

This study adopts five commonly used data mining methods to construct the prediction models, namely stepwise regression (SR), M5P, random forest (RF), support vector regression (SVR), and backpropagation neural network (BPN). In multiple regression analysis, SR is a method of selecting predictor variables and adding them to the regression equation. SR is a statistical method commonly used for predictions; it selects a few representative variables from multiple variables to construct prediction models. M5P method, proposed by Wang and Witten [57], is a modification of Quinlan [42]'s M5 tree algorithm that can predict continuous variables. It combines the conventional decision tree with linear regression functions at the node. Unlike other algorithms, M5P's advantage is that it applies to categorical and continuous variables and can handle missing values. RF, proposed by Breiman [13], is a combined learning algorithm based on decision tree classifiers. The classifier in a typical RF algorithm is CART (classification and regression tree). Multiple random variable samples select as the training dataset through bagging to reduce variation and help prevent overfitting. Many classification trees corresponding to the selected samples are constructed in the data training process to form RF. Finally, all the classification trees merge, and each classification votes on. The winning classification is selected based on the votes, and the final classification is determined. SVR, developed by Vapnik et al. [53], added ε - the insensitivity concept to the support vector machines (SVM) algorithm. SVR is a machine learning method widely used in prediction problems (Hong [25], Kim [29], Koike and Takagi [30]). This method uses training data to construct a regression equation. The goal is to minimize errors of the testing data inputted into the equation, achieving the purpose of prediction. The SVR model can determine the direction of each predictor variable and the standardized weight coefficients. If the direction is positive, the variable positively affects the prediction of the target variable. Otherwise, it has a negative effect. The larger the standardized weight coefficient, the more impact the variable has on the target variable prediction. BPN is an algorithm that imitates the biological neural network. It is one of the best examples of AI prediction methods and is widely used in various predictions (Zhang et al. [59]). Through error backpropagation, BPN's training method uses the gradient steepest descent to adjust the weights repeatedly through iterations. It is a process that minimizes the error between training data observed values and predicted outputs.

Although many existing studies have highlighted that deep learning techniques have a significant advantage over machine learning methods for operating with complex, high-dimensional, and extensive data, their disadvantages including the lack of variable selection ability, high computational costs, and massive data requirement have limited the use of deep learning in all fields (Singh et al. [49]). The main objective of this study is to find and discuss effective key elements for promoting HRM practices. In addition, the quantity and nature of the collected HRM data of this study are not suitable for deep learning modeling. Therefore, deep learning techniques are not considered in this study.

3.3. Proposed scheme

This study takes the human resource domain-driven approach and aims to solve problems in business practices. We identify critical variables through literature review, collect relevant data, and combine long-term dynamic information with data mining to identify essential hidden information and knowledge, thereby providing decision-making information for improving HRM practices. Figure 1 shows the proposed hybrid data mining scheme for job performance prediction. The scheme shown in the figure applies to the data mining model's development in all three years (Year One to Year Three). Only the target variables and predictor variables are different each year.

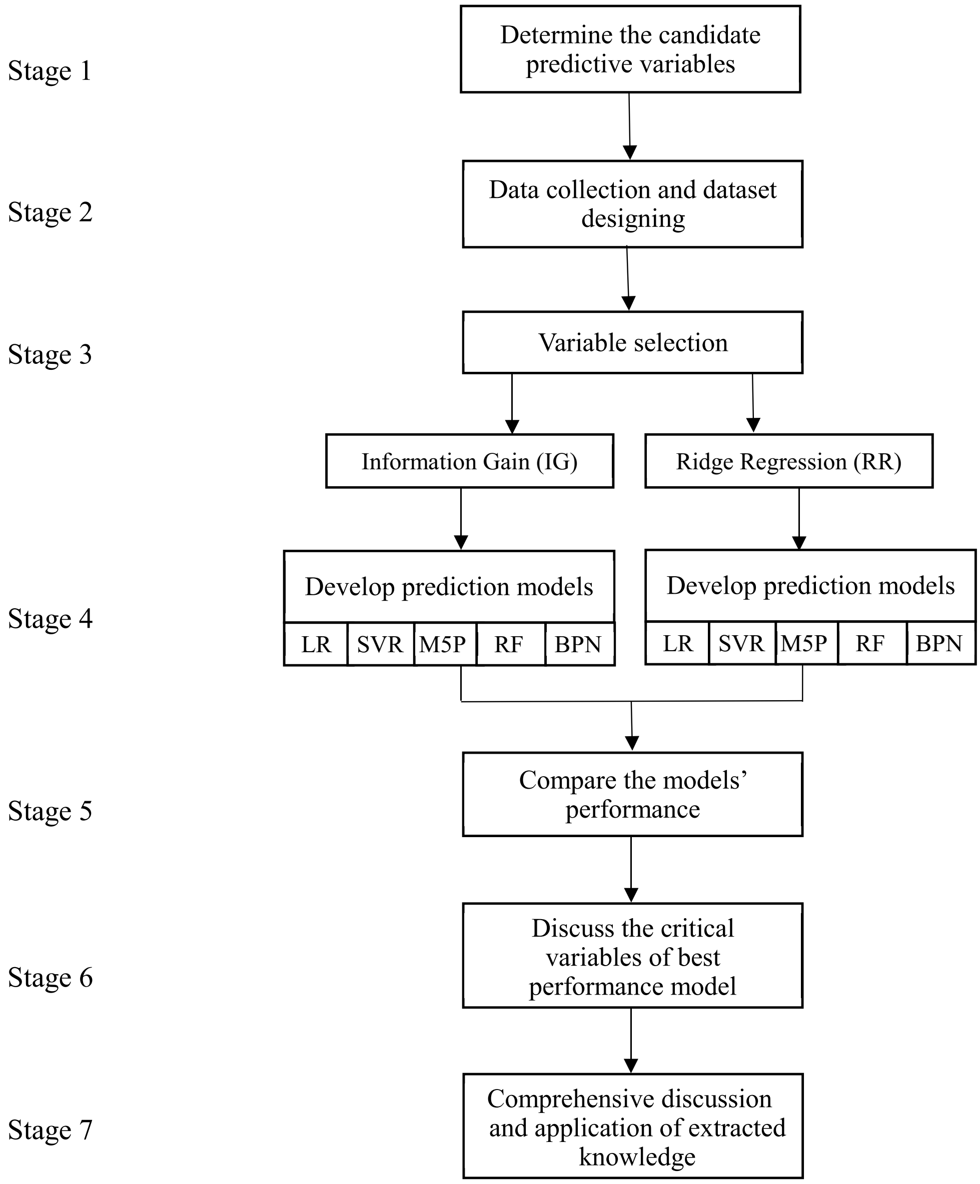

Figure 1: The proposed hybrid data mining scheme for job performance prediction.

As shown in Figure 1, stage 1 determines the candidate predictive variables. The variables are selected through literature review and suggested by the Chief Human Resources Officer of the case company and two professional scholars. Three experts determine that these variables are relevant and feasible for this study. The information used in constructing the models includes employees' personal data upon application, recruitment tests, work behavior, and performance feedback. There are a total of 37 variables selected in the study. Please refer to Table 1 for variable descriptions.

After determining the relevant predictor variables, the study enters stage 2: data

Table 1: List of the predictor variables and description.

Variables	**Description**
X1-Age	Age in performance year
X2-Gender	0-Female /1-Male
X3-EDU	Education, divided into three groups: master's and above, bachelor, high school and below
X4-School	Graduated school, divided into six tiers: N1:public level 1, N2:public level 2, P1:private level 1, P2:private level 2, F1:foreign school, OT:others)
X5-R-Depts	Major in industry-related departments: (0-No /1-Yes)
X6-Grade	Position level in the organization.
X7-IQ	Intelligence test scores.(0∼100)
X8-Active	It evaluates if a person can take quick actions or not.(0∼10)
X9-Continuative	It evaluates if a person has perseverance to finish changeless work. (0∼10)
X10-Commander	It evaluates if a person can know quickly his advantage in terms of his relationship with others and then affect others. (0∼10)
X11-Challenging	It evaluates if a person is willing to challenge new subjects or goals very often. (0∼10)
X12-Sympathy	It evaluates the extent of a person's empathy with others. (0∼10)
X13-Emotion	It evaluates a person's emotional stability, which means if his emotion is affected by the environment very often. (0∼10)
X14-Independent	It evaluates if a person can exert his creativity in his own way. (0∼10)
X15-Innovative	It evaluates the extent to which a person denies the current situation and has the desire to change it. (0∼10)
X16-Analytic	It evaluates if a person can often find the problems, analyze problems and make plans to solve them. (0∼10)
X17-Flexibility	It evaluates if a person can make judgments freely without being confined to some certain principles. (0∼10)
X18-Aesthesia	It evaluates the extent of a person's responsiveness and impact on outside stimulus. (0∼10)
X19-Deliberate	It evaluates if a person thinks carefully before doing anything and follows the plan step by step. (0∼10)
X20-Routine	Daily Routine/Processes capability(0∼10)
X21-People	People/Service capability(0∼10)
X22-Sales	Sales/Performance capability(0∼10)
X23-Planning	Strategic Planning capability(0∼10)
X24-Creative	Creative capability(0∼10)
X25-Leadership	Leadership potential(0∼10)
X26-Y0 Promote	Promoted after probation(0-No /1-Yes)
X27-AnuLea-times	Annual leave times during the year
X28-SickLea-times	Sick leave times during the year
X29-PerLea-times	Personal leave times during the year
X30-OT	Overtime hours during the year
X31-Y1-Performance	First year performance score(2∼8)
X32-Y1-Potential	First year potential score(2∼6)
X33-Y1 Promote	Promoted in the first year(0-No /1-Yes)
X34-Y2-Performance	Second year performance score(2∼8)
X35-Y2 Potential	Second year potential score(2∼6)
X36-Y2 Promote	Promoted in the second year(0-No /1-Yes)
X37-Y3 Performance	Third year performance score(2∼8)

collection and dataset designing. The data used in this study provides by the case company's human resources department after the end of each year. Regarding the characteristics of the sample, the ratio of males to females in year one is 27%: 73%, year two is 25%: 75% and year three is 26%: 74%. The ratio of the educational level—master, bachelor, senior high school, and below is 30%: 67%: 3% in year one, 33%: 64%: 3% in year two, and 34%: 62%: 4% in year three. The other variables descriptive statistics are shown in Table 2.

This study divides into three datasets—first-year, second-year, and third-year performance prediction—based on when the entry-level employees join the organization. This division is for subsequent data analysis of different time dimensions and observation of changes to critical variables in different periods. Providing accurate predictor variables and collecting data, this study divides predictor variables into three categories: fixed, annually-updated, and annual-new variables. Fixed predictor variables are not subject to change in the short run. They include personal information: X2-Gender, X3-EDU(education level), X4-School(graduation school classification), X5-R-Depts(whether majoring in industry-related disciplines), X7-IQ(intelligence test score), and personality test(X8-Active, X9-Continuative, X10-Commander, X11-Challenging, X12-Sympathy, X13-Emotion, X14-Independent, X15-Innovative, X16-Analytic, X17-Flexibility, X18-Aesthesia, X19-Deliberate, X20-Routine, X21-People, X22-Sales, X23-Planning, X24-Creative, X25-Leadership) upon recruitment and the promotion after probation variable(X26-Y0-Promote). Annually-updated predictor variables refer to the status of the variable in the performance year, including X1-Age, X6-Grade(position level), and work behavior: X27-AnuLea-times(annual leave times), X28-SickLea-times(sick leave times), X29-PerLea-times(personal leave times), and X30-OT(overtime hours). Annual-new predictor variables include performance, potential rating, and whether to promote in the previous year, including X31-Y1-Performance, X32-Y1-Potential, X33-Y1-Promote, X34-Y2-Performance, X35-Y2-Potential, X36-Y2-Promote), which are predictor variables for the following year. Therefore, the target variable in the first year is X31-Y1-Performance, and X1$\sim$X30 are predictor variables. The target variable in the second year is X34-Y2-Performance, and X1$\sim$X33 are predictor variables. The target variable in the third year is X37-Y3-Performance, and X1$\sim$X36 are predictor variables.

Stage 3 is variable selection. Variable selection is an essential step and determining factor in constructing prediction models. Reducing the use of irrelevant or redundant variables can speed up the algorithm's running time and improve the model's performance. This study adopts IG and RR two methods. Having selected different variable datasets via these two methods, we then use different algorithms to develop the model and determine the best composite model for each year.

Stage 4 is developing performance prediction models. After selecting the variables in stage 3, we use the selected predictor variables to construct job performance prediction models based on SR, M5P, RF, SVR, and BPN. At this stage, we construct 15 performance prediction models for each year. The two-stage performance prediction models generated by using IG as a variable selection method before each of the five perdition methods are represented by, respectively, IG-SR, IG-M5P, IG-RF, IG-SVR, and IG-BPN; the models constructed by using the RR variable selection method respectively term RR-SR, RR-M5P, RR-RF, RR-SVR, and RR-BPN. For evaluating the performance of the

Table 2: Descriptive statistics of year one to year three.

Variables	Year one				Year two				Year three			
	Min.	Max.	Mean	SD.	Min.	Max.	Mean	SD.	Min.	Max.	Mean	SD.
X1-Y1Age	23	54	28.89	5.13	24	53	29.99	5.07	25	54	31.35	5.41
X2-Gender	0	1	0.27	5.44	0	1	0.25	0.43	0	1	0.26	0.44
X5-R-Depts	0	1	0.28	0.45	0	1	0.29	0.46	0	1	0.29	0.45
X6-Grade	5	17	10.27	1.99	7	18	11.31	2.16	7	18	12.60	2.44
X7-IQ	24	90	56.90	11.19	26	90	56.81	11.10	26	90	56.72	11.59
X8-Active	2	10	7.46	1.42	2	10	7.48	1.49	2	10	7.37	1.49
X9-Continuative	3	10	8.32	1.56	3	10	8.32	1.55	3	10	8.32	1.51
X10-Commander	1	10	5.63	1.81	1	10	5.76	1.77	1	10	5.71	1.77
X11-Challenging	2	10	8.71	1.44	3	10	8.74	1.33	3	10	8.63	1.37
X12-Sympathy	1	10	5.76	1.96	1	10	5.83	1.97	2	10	5.83	2.05
X13-Emotion	2	10	8.51	1.56	3	10	8.49	1.56	3	10	8.46	1.61
X14-Independent	0	9	4.14	1.40	1	9	4.14	1.33	1	9	4.07	1.36
X15-Innovative	1	10	6.27	2.12	1	10	6.11	2.21	1	10	6.05	2.20
X16-Analytic	1	10	6.97	1.71	1	10	7.08	1.67	1	10	7.11	1.72
X17-Flexibility	2	10	8.19	1.50	2	10	8.09	1.44	3	10	7.99	1.45
X18-Aesthesia	2	9	5.17	1.31	2	9	5.22	1.27	2	8	5.22	1.22
X19-Deliberate	0	10	4.65	1.99	0	9	4.68	1.96	0	9	4.78	1.98
X20-Routine	0	7	3.37	1.38	0	7	3.22	1.36	1	6	3.21	1.35
X21-People	1	8	4.13	1.34	1	8	4.00	1.22	1	8	3.78	1.09
X22-Sales	1	10	5.53	1.25	1	10	5.56	1.25	1	9	5.52	1.21
X23-Planning	0	8	3.95	1.48	0	8	3.88	1.56	1	8	3.77	1.47
X24-Creative	0	6	2.65	1.27	0	6	2.72	1.25	0	6	2.82	1.27
X25-Leadership	0	8	4.80	1.47	0	8	4.74	1.46	0	8	4.61	1.43
X27-AnuLea-times	1	42	8.26	6.07	1	65	10.70	9.50	2	58	12.67	7.31
X28-SickLea-times	0	40	1.51	4.03	0	25	1.37	3.06	0	15	1.17	3.00
X29-PerLea-times	0	39	1.11	3.98	0	37	1.13	4.15	0	55	0.90	5.20
X30-OT	0	670	147.10	112.72	0	396	115.43	95.31	0	437.5	101.94	95.66
X31-Y1-Performance	2	8	5.54	1.48	3	8	5.61	1.47	3	8	5.68	1.48
X32-Y1-Potential					2	6	4.75	1.06	2	6	4.82	1.06
X33-Y1 Promote					0	1	0.37	0.49	0	1	0.36	0.48
X34-Y2-Performance					2	8	6.13	1.58	3	8	6.26	1.49
X35-Y2 Potential									2	6	5.05	0.99
X36-Y2 Promote									0	1	0.55	0.50
X37-Y3-Performance									3	8	6.36	1.41

ten two-stage prediction models, the single models without using a variable selection method construct and term SR, M5P, RF, SVR, and BPN, respectively.

Stage 5 is comparing model performance. This study uses two indicators to evaluate model performance: mean absolute error (MAE) and root mean squared error (RMSE). MAE is the mean absolute difference between the test samples' predictive and observed values, and RMSE is the square root of the mean squared deviation between predictive and observed values. The more significant the difference between predictive and observed values, the heavier the weight. Both indicators use how close the model's predictive values are to the observed values to show how well the model fits the data. The smaller the value, the better the model.

Stage 6 discusses critical variables related to the prediction performance. According to the 15 performance prediction models constructed for each year in stage 4 and the comparison using the two performance indicators in stage 5, we select the best performance prediction model for each year. We discuss these critical variables of the models based on the materiality principle. This stage aims to find the most influential variables in each year and discuss their implications for subsequent comprehensive discussion and knowledge extraction from year one to year three.

Stage 7 is the three-year comprehensive discussion and suggestions for applying the knowledge extracted. By comparing and discussing critical variables in years one to three, we identify influential critical variables and their changes in different stages. We extract knowledge and provide suggestions for the application of the management of HRM practices.

This study uses the WEKA 3.8.3 version as the data mining software suite. WEKA (Witten et al. [58]) is a software suite used for machine learning and data mining, and it is widely used in academic and industrial research. This study uses the ten folds cross-validation to construct the models. All parameters use the default setting of WEKA 3.8.3's algorithm, with no additional parameter performance tuning.

4. Empirical Results

The data used for this study was sourced from a leading fast fashion manufacturing company in Taiwan. Its clients are well-known American, European, and Japanese brands, department stores, and distributors. The case company performed very well in both financial and non-financial indicators from 2014 to 2018. In the EPS part of financial indicators, it was in the top 10% of 55 companies in the same industry. In non-financial indicators- Corporate Governance Evaluation, it was in the top 20% of all listed companies, which shows that the case company has a substantial competitive advantage in the global market and is a suitable research target.

This study collects data on entry-level employees who joined the company of the case study between 2014 to 2018. According to employee tenure, we categorize the data into three datasets—performance in the first, second, and third years. Each year's samples are 391, 227, and 164, respectively.

4.1 Variable selection results

According to the proposed framework, we first select critical variables using the IG and RR methods, respectively. Tables 3 to 5 respectively show the selected critical variables in year one to year three yielded by the two variable selection methods—the variables using the IG method selected in the first year with six variables, including X5-R-Depts. Eleven variables, including X1-Age, are selected using the RR method. Through comparison, we find that five of the six variables by the IG method are consistent with

Table 3: Variable selection results using IG and RR methods-year one.

Methods	Selected variables
IG	X5-R-Depts
	X6-Grade
	X7-IQ
	X21-People
	X26-Y0 Promote
	X30-OT
IG	X1-Age
	X2-Gender
	X4-School
	X5-R-Depts
	X6-Grade
	X7-IQ
	X17-Flexibility
	X19-Deliberate
	X21-People
	X24-Creative
	X30-OT

Table 4: Variable selection results using IG and RR methods-year two.

Methods	Selected variables
IG	X4-School
	X7-IQ
	X13-Emotion
	X23-Planning
	X25-Leadership
	X26-Y0 Promote
	X30-OT
	X31-Y1-Performance
	X32-Y1-Potential
RR	X1-Age
	X6-Grade
	X7-IQ
	X13-Emotion
	X18-Aesthesia
	X25-Leadership
	X30-OT
	X31-Y1-Performan
	X32-Y1-Potential
	X33-Y1 Promote

Table 5: Variable selection results using IG and RR methods- year three.

Methods	Selected variables
IG	X2-Gender
	X3-EDU
	X4-School
	X6-Grade
	X7-IQ
	X9-Continuative
	X15-Innovative
	X26-Y0 Promote
	X28-SickLea-times
	X29-PerLea-times
	X32-Y1-Potential
	X33-Y1-Promote
	X34-Y2-Performance
	X35-Y2-Potential
	X36-Y2-Promote
RR	X1-Age
	X3-EDU
	X5-R-Depts
	X6-Grade
	X7-IQ
	X9-Continuative
	X13-Emotion
	X28-SickLea-times
	X32-Y1-Potential
	X34-Y2-Performance

the RR method, namely X5-R-Depts, X6-Grade, X7-IQ, X21-People, and X30-OT. Only one variable, X26-Y0 Promote, is not found in the variables by the RR method. Six of the eleven variables by the RR method are not found in the variables by the IG method, namely X1-Age, X2-Gender, X4-School, X17-Flexibility, X19-Deliberate, and X24-Creative.

In the second year, nine variables, including X4-School, are selected using the IG method. Ten variables, including X1-Age, are selected using the RR method. Through comparison, we find that six variables are consistent, and seven are not. The seven are the three variables by the IG method, X4-School, X23-Planning, and X26-Y0 Promote; the four variables by the RR method, X1-Age, X6-Grade, X18-Aesthesia, and X33-Y1-Promote.

Fifteen variables, including X2-Gender, are selected in the third year using the IG method. Ten variables, including X1-Age, are selected using the RR method. Through comparison, we find that seven variables are consistent, and eleven are not. The eleven are the eight variables by the IG method, X2-Gender, X4-School, X15-Innovative, X26-Y0 Promote, X29-PerLea-times, X33-Y1-Promote, X35-Y2-Potential, and X36-Y2-Promote; the three variables by the RR method, X1-Age, X5-R-Depts, and X13-Emotion.

In addition, the variable data by the IG and RR methods show that each method goes by a different selection logic. Therefore, there is a considerable gap between the variables selected, which may have a critical impact on the performance of subsequent prediction models.

4.2. Model performance evaluation

According to the process of the proposed scheme mentioned in section 3, we construct 15 performance prediction models for each year. The two indicators, MAE and RMSE, are used to evaluate model performance and determine the best performance prediction model for each year.

Table 6: Performance comparison of 15 prediction models for year one to year three.

	Year One		**Year Two**		**Year Three**	
Methods	**MAE**	**RMSE**	**MAE**	**RMSE**	**MAE**	**RMSE**
SR	1.1306	1.3847	1.1363	1.3884	1.0371	1.3067
M5P	1.1440	1.4170	1.1537	1.4216	1.0385	1.3081
RF	1.2028	1.4187	1.1524	1.3415	1.0636	1.2981
SVR	1.1233	1.4069	1.2049	1.5053	1.0863	1.3663
BPN	1.8096	2.2955	1.6281	2.0958	1.4527	1.8517
IG-SR	1.1196	1.3528	1.0903	1.3225	1.0390	1.2877
IG-M5P	1.1196	1.3528	1.0891	1.3213	1.0328	1.2850
IG-RF	1.1273	1.3587	1.1395	1.3482	1.0432	1.2718
IG-SVR	1.1101	1.3759	1.0964	1.3542	1.0477	1.3326
IG-BPN	1.4592	1.8056	1.7364	2.2341	2.0388	2.5942
RR-SR	1.1145	1.3576	1.0669	1.3131	0.9476	1.1768
RR-M5P	1.1293	1.3911	1.0997	1.3401	0.9669	1.1947
RR-RF	1.1798	1.3960	1.1262	1.3404	0.9861	1.2289
RR-SVR	**1.0691**	**1.3446**	**1.0457**	**1.2902**	**0.9022**	**1.1484**
RR-BPN	1.5902	2.1182	1.5048	2.0053	1.5221	2.0436

Table 6 shows the prediction performance of the 15 models from year one to year three. It can be seen from the table that the RR-SVR model performed best and most stable in both MAE and RMSE metrics in three years. Therefore, based on the RR-SVR model, the critical variables of each year selected will be discussed.

5. Discussion

In the following, first, we discuss the critical variables of each year. Then, we discuss the critical variables in all three years to find the influential ones and the changes and implications of variables throughout the years.

It is worth noting that this study uses the materiality principle to select critical variables to meet the demands of management practices. Thus, to avoid losing focus due to excess variables, we use the absolute value of the standardized weight coefficient of each variable provided by the RR-SVR model to rank the selected variables in decreasing order. The meanings of the first five variables with larger weights are discussed. Table 7 shows the variables' ranking, direction, and weights generated by the RR-SVR performance prediction model for the three years.

Table 7: The variables' ranking, direction, and weights generated by the RR-SVR performance prediction model for the three years.

Ranking	Year One			Year Two			Year Three		
	Direction	Weights	Variables	Direction	Weights	Variables	Direction	Weights	Variables
1	+	0.8434	X6-Grade	+	0.3125	X30-OT	-	0.3589	X9-Continuative
2	+	0.4764	X30-OT	+	0.3008	X31-Y1-Performance	+	0.3570	X6-Grade
3	-	0.3737	X1-Age	+	0.2312	X6-Grade	-	0.2741	X1-Age
4	-	0.2843	X21-People	-	0.2119	X1-Age	+	0.2621	X34-Y2-Performance
5	+	0.2017	X7-IQ	+	0.1877	X7-IQ	+	0.2010	X7-IQ
6	+	0.1729	X19-Deliberate	+	0.1829	X18-Aesthesia	+	0.1664	X13-Emotion
7	-	0.1495	X24-Creative	+	0.1451	X32-Y1-Potential	+	0.1290	X3-EDU=Senior
8	+	0.1403	X17-Flexibility	-	0.1409	X25-Leadership	+	0.1154	X32-Y1-Potential
9	+	0.0800	X2-Gender=F	+	0.1313	X13-Emotion	-	0.1027	X3-EDU=Bachelor
10	+	0.0530	X4-School=N1	-	0.0632	X33-Y1 Promote=Y	-	0.0867	X5-R-Depts=N

5.1. The critical variables of year one

Table 5 shows that the first five critical variables to predict job performance in year one are X6-Grade, X30-OT, X1-Age, X21-People, and X7-IQ, respectively. First, the X6-Grade(position level) is the employee's appointment by the organization based on the employee's experience and ability. The higher the grade, the higher the organization's appraisal of the new employee's experience and ability. X6-Grade has the highest weight coefficient and a positive correlation, showing that relevant work experience significantly impacts performance in the first year. This result is consistent with Quińones et al. [43]'s meta-analysis: Work experience and job performance positively correlated. Therefore, work experience is essential for new employees to perform well quickly.

The second variable is X30-OT (overtime hours), and overtime positively correlates with job engagement. This result is consistent with Schaufeli et al. [46]'s study and Shimazu et al. [48]'s study and also consistent with the social exchange theory (Homans [24]).

The X1-Age variable negatively correlates with performance and is ranked third. This study adopts appraisal by supervisors to evaluate performance. The result is consistent with Waldman and Avolio [54]'s meta-analysis and Saks and Waldman [45]'s study. It also shows that when selecting employees based on future development, it is more objective to consider experience and age.

The variable affecting performance that ranks fourth is X21-People (interpersonal skills). It shows a negative correlation to job performance. This skill is suitable for highly interactive positions like reception, customer service, recruitment manager, tour guide, Etc. The demand for such positions is relatively low in the fast fashion industry. Most tasks are core business-related such as order development, procurement, production, and handling abnormal events. Daily operations are fast-paced, and employees must respond quickly to market opportunities. Competence in this skill may affect the pace of core business. Therefore, it is reasonable that it shows a negative correlation.

X7-IQ (intelligence test score) affects performance that ranks fifth and positively correlates to job performance. The intelligence test evaluates one's ability to think, learn, and adapt to the environment. Although intelligence test score is not one of the top three variables affecting performance, it still plays an important role. The result is consistent with McClelland [34]'s suggestion: Intelligence may be related to job performance, but there is no absolute causality. It should consider other variables (e.g., gender, education, personality, Etc.).

5.2. The critical variables of year two

It also can be seen from Table 5 that the first five critical variables to predict job performance for year two are X30-OT, X32-Y1 Performance, X6-Grade, X1-Age, and X7-IQ, respectively. Four of the variables are consistent with the variables in the first year. Please refer to the explanation in the first year for the meaning of relevant variables. The new variable is X32-Y1-Performance (performance in the first year), which ranks second. Its direction indicates that previous job performance considerably impacts current job performance.

The overall ranking and weight changes have significant implications. The first ranking of the prediction models changed from X6-Grade in the first year to X30-OT (overtime hours) in the second year. X6-Grade that previously ranked first fell to the third-ranking. Also, X32-Y1-Performance becomes the second most influential variable. It shows that work experience no longer plays the most critical variable, and its influence on performance decreases. The most influential variables are job engagement (overtime) and previous performance (Y1).

Besides X30-OT and X6-Grade, X1-Age and X7-IQ remain in the top five influential variables. It shows that the year two model sustains the year one model, and the variables are predictively stable. Apply these findings to the company's practice. Suppose the new employees have high engagement (overtime), relevant work experience, are relatively young, have high intelligence test scores, and perform well in the first year. In that case, the new employees will be more likely to perform well in the second year.

5.3. The critical variables of year three

The first five critical variables to predict job performance for year three are X9-Continuative, X6-Grade, X1-Age, X34-Y2 Performance (performance in the second year), and X7-IQ, which also can observe in Table 5. Four of the variables are consistent with the variables in the second year. They are X6-Grade, X1-Age, X34-Y2-Performance (note that it is the performance in the previous year, the same meaning as X31-Y1-Performance in year two), and X7-IQ. However, the ranking first changes to X9-Continuative (persistence) and negatively correlates to job performance. Persistence is the degree of perseverance in completing continuous and fixed tasks. High persistence also means low flexibility. As the industry characteristics described above, the fast fashion industry is constantly changing and involves many subjective decisions. At this point, employees have joined the company for three years. They should be able to work and solve problems independently. The inability to solve problems flexibly will lead to more significant problems. Therefore, high persistence is negatively correlated with performance.

Ranking 2 is X6-Grade. In general, the higher-grade employees this year were promoted from lower grades in the first and second years. When combined with X34-Y2-Performance, which ranks fourth, we find that employees who perform well and have been promoted are more likely to perform well in the future, especially compared to younger employees (X1-Age).

X7-IQ has been among the top five influential variables for three consecutive years. It shows that it is an influential critical variable in the new employees' short to medium tenure (one to three years), consistent with McClelland [34]'s finding.

5.4. Comprehensive discussion of critical variables of all three years

From the above discussion of critical variables in year one to year three, we find many recurring variables, namely X1-Age, X6-Grade, X7-IQ, and performance in the previous year (X31-Y1 Performance, X34-Y2 Performance). The finding shows that these variables have a long-term, stable predictive effect on performance. However, the same variables have different implications in different years. For example, X6-Grade has the most influence on performance in the first year; showing relevant work experience of the new employees is essential. As the employees matured and the grade structure changed in the second year, X30-OT and performance in the previous year became more important. X30-OT is a critical variable affecting performance in the first and second years, but it is no longer a variable affecting performance in the third year. It shows that the new employees mainly were assistants in the early stage, and supervisors may have focused more on attitude and engagement for performance appraisal. However, when an employee has been in the company for three years, they should be able to work independently and be fully responsible for their own responsibilities. At this time, the focus of the supervisor's performance evaluation on his work is not the number of overtime hours but the degree to which the employee has fulfilled his responsibilities. Also, when analyzing X6-Grade, X1-Age, and X7-IQ together, we can reasonably say that if the company urgently needs short-term labor, those with relevant work experience should be able to deliver quickly. However, in terms of mid- to long-term development, younger people who have met the intelligence test requirement and have relevant professional experience are the better fit. We presume that employees who perform well in the first year and are highly engaged (X30-OT) are more likely to perform well in the future. The company should invest more resources in these employees to maximize their performance.

Variables related to personality and ability include X10-Continuative, ranking first in the third year, and X21-People, ranking fourth in the first year. We suggest the case company discuss in depth the key personality and abilities suitable for its tasks and apply them to the management and performance appraisal of new employees. If doing so, the case company can achieve person-organization fit and increase employees' organizational commitment and task performance (Kristof-Brown et al. [31]). Preventing competent employees from leaving due to managerial issues or short-term difficulty adapting will affect the company's long-term cultivation of employees.

Lastly, the variable selection and changes of critical variables in three years show that the growth of new employees is a dynamic process. The single static prediction model cannot effectively explain the changes in this dynamic process, showing that previous static prediction models are insufficient. The results of this study should compensate for the shortcomings of these models. This finding is also consistent with the viewpoint of Bieńkowska and Tworek [11] that employees' dynamic capabilities (EDC) are important factors influencing job performance. In this study, some critical variables are essential from year one to year three (grade, age, intelligence, and performance in the previous year). It means that the company should consider these variables first upon hiring. Some of the variables changed within the three years, showing the predictive variables of performance change with the employees' stage of development. The company should pay more attention to this issue and adjust its short-, medium-, and long-term labor demand and development plans. The results can be used for recruitment and employee development, training, compensation, and resource allocation.

6. Conclusion

This study combines the two disciplines of human resources and data mining and proposes a hybrid data mining scheme for job performance prediction. We collect three years of data and adopt two variable selection methods and five data mining techniques to construct job performance prediction models of the human resource domain for the fast fashion industry. We identify the critical variables affecting the job performance of new employees through a discussion of the prediction models in three years. We find that relevant critical variables are dynamic and have a different impact. Also, the knowledge extracted in this study is helpful for practice. Unlike static research, the company can continuously use this method in the actual operation in the future. Through the hybrid data mining scheme, we can understand the changes in essential variables affecting employees' job performance and adjust human resources strategies to response the changes in the external business environment, increase the value of human resources, and help the company to build sustainable competitive advantages.

This study has some limitations. First, this study is a case study centered on the evaluation process of model development and discussion of the knowledge extracted, and the results may not apply to other industries or companies. The second limitation concerns data comprehensiveness. We could not collect all the data on performance. If future studies can achieve data comprehensiveness, the prediction models and implications for management should be more valuable. Lastly, this study does not use all the variable selection methods and algorithms applicable to model development. Therefore, the results of this study are the best model among the five data mining techniques used and cannot represent the best model among all the mining techniques. Future studies can evaluate these effects, which should increase the value of the research.

References

[1] Abelln, J. and Castellano, J. G. (2017). *Improving the Naive Bayes Classifier via a Quick Variable Selection Method Using Maximum of Entropy*, Entropy, Vol.19, 247.

[2] Al-Radaideh, Q. A., and Al Nagi, E. (2012). *Using Data Mining Techniques to Build a Classification Model for Predicting Employees Performance*, International Journal of Advanced Computer Science and Applications, Vol.3.

[3] Alam, T. M., Mushtaq, M., Shaukat, K., Hameed, I. A., Umer Sarwar, M. and Luo, S. (2021). *A Novel Method for Performance Measurement of Public Educational Institutions Using Machine Learning Models*, Applied Sciences, Vol.11, 9296.

[4] Aluko, R. O., Daniel, E. I., Shamsideen Oshodi, O., Aigbavboa, C. O. and Abisuga, A. O. (2018). *Towards Reliable Prediction of Academic Performance of Architecture Students Using Data Mining Techniques*, Journal of Engineering, Design and Technology, Vol.16, 385-397.

[5] Arefeen, M. A., Nimi, S. T., Rahman, M. S., Arshad, S. H., Holloway, J. W. and Rezwan, F. I. (2020). *Prediction of Lung Function in Adolescence Using Epigenetic Aging: A Machine Learning Approach*, Methods and Protocols, Vol.3, 77.

[6] Arfaee, M., Bahari, A. and Khalilzadeh, M. (2022). *A Novel Prediction Model for Educational Planning of Human Resources With Data Mining Approach: A National Tax Administration Case Study*, Education and Information Technologies, Vol.27, 2209-2239.

[7] Bach, M. P., Šimić, N. and Merkač, M. (2013). *Forecasting Employees' Success at Work in Banking: Could Psychological Testing Be Used as the Crystal Ball?*, Napovedovanje uspešnosti dela zaposlenih v bančništvu: je mogoče psihološko testiranje uporabljati kot kristalno kroglo?, Vol.11, 283-299.

[8] Backhaus, K. B., Stone, B. A., and Heiner, K. (2002). *Exploringthe Relationship Between Corporate Social Performance and Employer Attractiveness*, Business Society, Vol.41, 292-318.

[9] Barry, A., Li, W., Becerra, J. A. and Gilabert, P. L. (2021). *Comparison of Feature Selection Techniques for Power Amplifier Behavioral Modeling and Digital Predistortion Linearization*, Sensors, Vol.21, 5772.

[10] Becker, B. and Gerhart, B. (1996). *The Impact of Human Resource Management on Organizational Performance: Progress and Prospects*, Academy of Management Journal, Vol.39, 779-801.

[11] Bieńkowska, A., and Tworek, K. (2020). *Job Performance Model Based on Employees' Dynamic Capabilities (EDC)*, Sustainability, Vol.12, 2250.

[12] Borman, W. C. and Motowidlo, S. M. (1993). Expanding The Criterion Domain to Include Elements of Contextual Performance, Jossey-Bass, In N. Schmitt, W. C. Borman, and Associates, Personnel Selection in Organization., San Francisco.

[13] Breiman, L. (2001). *Random Forests*, Machine Learning, Vol.45, 5-32.

[14] Campbell, J. P. (1990). *Modeling The Performance Prediction Problem in Industrial and Organizational Psychology*, In Handbook of Industrial and Organizational Psychology, Vol.1, 2nd ed., pp 687-732, Consulting Psychologists Press, Palo Alto, CA, US.

[15] Christopher, M., Lowson, R. and Peck, H. (2004). *Creating Agile Supply Chains in The Fashion Industry*, International Journal of Retail Distribution Management, Vol.32, 367-376

[16] Delaney, J. T. and Huselid, M. A. (1996). *The Impact of Human Resource Management Practices on Perceptions of Organizational Performance*, Academy of Management Journal, Vol.39, 949-969.

[17] Delgado-Gómez, D., Aguado, D., Lopez-Castroman, J., Santacruz, C. and Arts-Rodriguez, A. (2011). *Improving Sale Performance Prediction Using Support Vector Machines*, Expert Systems with Applications, Vol.38, 5129-5132.

[18] Festinger, L. (1957). A theory of Cognitive Dissonance, Vol.2, Stanford university press.

[19] Festinger, L., and Aronson, E. (1960). The Arousal and Reduction of Dissonance in Social Contexts, Group dynamics, 214-231.

[20] Friedman, J., Hastie, T. and Tibshirani, R. (2010). *Regularization Paths for Generalized Linear Models via Coordinate Descent*, J. Stat. Softw, Vol.33, 1-22.

[21] Goto, K. and Endo, T. (2014). *Upgrading, Relocating, Informalising? Local Strategies in The Era of Globalisation: The Thai Garment Industry*, Journal of Contemporary Asia, Vol.44, 1-18.

[22] Hastie, T., Tibshirani, R. and Wainwright, M. (2015). Statistical Learning with Sparsity: The Lasso and Generalizations, CRC Press, Boca Raton, FL, USA.

[23] Hoerl, A. E. and Kennard, R. W. (1970). *Ridge Regression: Biased Estimation for Nonorthogonal Problems*, Technometrics, Vol.12, 55-67.

[24] Homans, G. C. (1958). *Social Behavior as Exchange*, American journal of sociology, Vol.63, 597-606.

[25] Hong, W.-C. (2011). *Traffic Flow Forecasting by Seasonal SVR with Chaotic Simulated Annealing Algorithm*, Neurocomputing, Vol.74, 2096-2107.

[26] Honglei, Z. (2009). *Fuzzy Evaluation on The Performance of Human Resources Management of Commercial Banks Based on Improved Algorithm*, In 2009 2nd International Conference on Power Electronics and Intelligent Transportation System (PEITS), 214-218.

[27] Judiesch, M. K., and Lyness, K. S. (1999). *Left Behind? The Impact of Leaves of Absence on Managers' Career Success*, Academy of Management Journal, Vol.42, 641-651.

[28] Kim, J., Lee, J. and Park, M. (2022). *Identification of Smartwatch-Collected Lifelog Variables Affecting Body Mass Index in Middle-Aged People Using Regression Machine Learning Algorithms and SHapley Additive Explanations*, Applied Sciences, Vol.12, 3819.

[29] Kim, K.-j. (2003). *Financial Time Series Forecasting Using Support Vector Machines*, Neurocomputing, Vol.55, 307-319.

[30] Koike, A. and Takagi, T. (2004). *Prediction of Protein-Protein Interaction Sites Using Support Vector Machines*, Protein Engineering, Design and Selection, Vol.17, 165-173.

[31] Kristof-Brown, A. L., Zimmerman, R. D. and Johnson, E. C. (2005). *Consequences of Individuals' Fit at Work: A Meta-Analysis of Person-Job*, Person-Organization, Person-Group, and Person-Supervisor Fit., Personnel Psychology, Vol.58, 281-342.

[32] Kwon, S., Lee, S., and Na, O. (2017). *Tuning Parameter Selection for The Adaptive LASSO in The Autoregressive Model*, Journal of the Korean Statistical Society, Vol.46, 285-297.

[33] Liu, C.-H., Tsai, C.-F., Sue, K.-L. and Huang, M.-W. (2020). *The Feature Selection Effect on Missing Value Imputation of Medical Datasets*, Applied Sciences, Vol.10, 2344.

[34] McClelland, D. C. (1993). *Intelligence Is Not the Best Predictor of Job Performance*, Current Directions in Psychological Science, Vol.2, 5-6.

[35] Murphy, K. R. and Cleveland, J. N. (1995). Understanding Performance Apraisal: Social, Organizational, and Goal-Based Perspectives, Sage.

[36] Nevis, E. C., DiBella, A. J., and Gould, J. M. (1995). *Understanding Organizations as Learning Systems*, Sloan Management Review, Vol.36, 73-85.

[37] Ng, T. W. H. and Feldman, D. C. (2009). *How Broadly Does Education Contribute To Job Performance?*, Personnel Psychology, Vol.62, 89-134.

[38] Noe, R. A., Hollenbeck, J. R., Gerhart, B. and Wright, P. M. (2017). Human Resource Management: Gaining a competitive advantage, McGraw-Hill Education New York, NY.

[39] Ogutu, J. O., Schulz-Streeck, T. and Piepho, H.-P. (2012). *Genomic Selection Using Regularized Linear Regression Models: Ridge Regression, Lasso, Elastic Net And Their Extensions*, BMC Proceedings, Vol.6, S10.

[40] Pejić Bach, M., Šimić, N., and Merkač, M. (2013). *Forecasting Employees' Success at Work in Banking: Could Psychological Testing Be Used as the Crystal Ball*?, Managing Global Transitions, Vol.11, 283-299.

[41] Peterson, D. K. (2004). *The Relationship Between Perceptions of Corporate Citizenship and Organizational Commitment*, Business Society, Vol.43, 296-319.

[42] Quinlan, J. R. (1992). Learning with Continuous Classes, In 5th Australian joint conference on artificial intelligence, 343-348, World Scientific, Singapore: World Scientific.

[43] Quiónes, M. A., Ford, J. K. and Teachout, M. S. (1995). *The Relationship Between Work Experience And Job Performance: A Conceptual And Meta-Analytic Review*, Personnel Psychology, Vol.48, 887-910.

[44] Ree, M. J., Earles, J. A., and Teachout, M. S. (1992). General Cognitive Ability Predicts Job Performance, ARMSTRONG LAB BROOKS AFB TX.

[45] Saks, A. M., and Waldman, D. A. (1998). *The Relationship between Age and Job Performance Evaluations for Entry-Level Professionals*, Journal of Organizational Behavior, Vol.19, 409-419.

[46] Schaufeli, W. B., Taris, T. W. and Van Rhenen, W. (2008). *Workaholism, Burnout, and Work Engagement: Three of a Kind or Three Different Kinds of Employee Well-being*?, Applied Psychology, Vol.57, 173-203.

[47] Schuler, R. S., and Jackson, S. E. (1987). *Linking Competitive Strategies with Human Resource Management Practices*, Academy of Management Perspectives, Vol.1, 207-219.

[48] Shimazu, A., Schaufeli, W. B., Kamiyama, K. and Kawakami, N. (2015). *Workaholism vs. Work Engagement: the Two Different Predictors of Future Well-being and Performance*, International Journal of Behavioral Medicine, Vol.22, 18-23.

[49] Singh, A., Sengupta, S. and Lakshminarayanan, V. (2020). *Explainable Deep Learning Models in Medical Image Analysis*, Journal of Imaging, Vol.6, 52.

[50] Strohmeier, S. and Piazza, F. (2013). *Domain Driven Data Mining in Human Resource Management: A Review of Current Research*, Expert Systems with Applications, Vol.40, 2410-2420.

[51] Teece, D. J., Pisano, G. and Shuen, A. (1997). *Dynamic Capabilities and Strategic Management*, Strategic Management Journal, Vol.18, 509-533.

[52] Valle, M. A., Varas, S. and Ruz, G. A. (2012). *Job Performance Prediction in A Call Center Using A Naive Bayes Classifier*, Expert Systems with Applications, Vol.39, 9939-9945.

[53] Vapnik, V., Golowich, S. E. and Smola, A. J. (1997). *Support Vector Method for Function Approximation, Regression Estimation and Signal Processing*, Advances in Neural Information Processing Systems, Vol.9, 281-287.

[54] Waldman, D. A. and Avolio, B. J. (1986). *A Meta-Analysis of Age Differences in Job Performance*, Journal of Applied Psychology, Vol.71, 33.

[55] Waldman, D. A. and Spangler, W. D. (1989). *Putting Together the Pieces: A Closer Look at the Determinants of Job Performance*, Human Performance, Vol.2, 29-59.

[56] Wang, K.-Y. and Shun, H.-Y. (2016). *Applying Back Propagation Neural Networks in the Prediction of Management Associate Work Retention for Small and Medium Enterprises*, Universal Journal of Management, Vol.4, 223-227.

[57] Wang, Y. and Witten, I. H. (1996). Induction of Model Trees for Predicting Continuous Classes. (Working paper 96/23), Hamilton, New Zealand: University of Waikato, Department of Computer Science.

[58] Witten, I., Frank, E. and Hall, M. (2011). Chapter 10-Introduction to Weka. Data Mining: Practical Machine Learning Tools and Techniques, Boston: Morgan Kaufmann.

[59] Zhang, G., Eddy Patuwo, B. and Y. Hu, M. (1998). *Forecasting with Artificial Neural Networks: The State of The Art*, International Journal of Forecasting, Vol.14, 35-62.

Graduate Institute of Business Administration, Fu Jen Catholic University, Taiwan, ROC.

E-mail: 407088136@mail.fju.edu.tw

Major area (s): Organizational behavior and human resource management.

Department of Business Administration, Fu Jen Catholic University, Taiwan, ROC.

E-mail: 051507@mail.fju.edu.tw

Major area (s): Organizational behavior and human resource management.

Graduate Institute of Business Administration, Fu Jen Catholic University, Taiwan, ROC.

E-mail: 059099@mail.fju.edu.tw (Corresponding Author)

Major area(s): Data mining, machine learning, medical/healthcare informatics,quality management.

(Received September 2022; accepted November 2022)