ORIGINAL ARTICLE

Using data mining technique to perform the performance assessment of lean service

Liu Ming-Te · Mei Albert Kuo-Chung ·
Wen-Tsao Pan

Abstract Lean production and service means to improve the production and service management within an enterprise. Through several management techniques, the waste, the redundant things, and the expenses without added values can be cleared, and the production and service within an enterprise will become smoother, that is, the competitiveness of an enterprise will be enhanced. In this study, the data collected from experiment carried out in Toyota lean service simulation laboratory of China University of Technology are used for analysis. Data mining technique is used to investigate whether the result of lean production and service taken by enterprises can enhance the entire performance of production and service. In this study, Grey relational analysis is performed first and is used to judge whether the data collected in the experiment using lean production and service can enhance performance; then, clustering method is used to classify experimental data into two clusters based on service attitude and dish-serving efficiency; finally, three data mining techniques of Genetic Programming (GP), Back-propagation Artificial Neural Network and logistic regression are used to set up, respectively, lean service performance model and Employee Characteristic Analysis model. From the analysis result, it is shown that the result of lean production and service can indeed enhance the performance of entire production and service; and among the three data mining techniques, GP model has the best classification and forecast capability.

1 Preface

In the globalization trend of enterprise nowadays, corporate operation has faced changeful impact. To survive, the enterprise has to adapt itself to the environment and change its business operation direction as well as promote business operation reform and innovative concept. However, during the first oil crisis, all the industries around the world are affected, the global market has weak demand, and the production cost is greatly enhanced; the economic recession has made Japan to reach almost a zero growth rate. Until now, the enterprise operation of many enterprises is still in dangerous state, and there is only one special case, that is, TOYOTA company of Japan, which is almost not affected by bad global environment; although the profit is slightly reduced, yet as compared to other enterprises, the profit is still large; therefore, Toyota's management has become scientific management paradigm well recognized around the world. All the global enterprises are attracted to Toyota's Production System (TPS) and try to study it and find a way to introduce it into their enterprises effectively. From the empirical study of Womack and Jones [1], it was pointed out that its success key is the so-called lean

L. Ming-Te · M. A. Kuo-Chung (✉)
Department of Business Administration,
China University of Technology, No. 56, Sec. 3,
Xinglong Rd., Wenshan District,
116 Taipei City, Taiwan, ROC
e-mail: mei@cute.edu.tw

L. Ming-Te
e-mail: mtliu@cute.edu.tw

W.-T. Pan
Department of Information Management,
Oriental Institute of Technology, 3F., No. 12,
Lane 271, Longjiang Rd., Jhongshan District,
104 Taipei City, Taiwan, ROC
e-mail: teacherp0162@yahoo.com.tw

production. Today, there are many large enterprises around the world that develop, from TPS, many lean production or service systems that are suitable to be applied in their own companies. However, currently, many researchers have been widely studying the lean production or service systems problem [2–4].

The Department of Business Administration of China University of Technology had received a support of "The research and training center of lean enterprise" project from "Union project of academy and research institute" of Department of Industrial Technology of the Ministry of Economic Affairs (MOEA) of Taiwan to form Toyota's lean service simulation laboratory. In the laboratory, the preparation of Taiwan's delicious foods is taken as an example to perform lean production and service simulation experiment, and Taiwan's delicious foods is as shown in Fig. 1. The experiment can be divided into three stages, the first stage is that the employees, under chaotic condition, perform the preparation of delicious foods, the second stage is the improvement of the preparation flow of delicious foods, and the third stage is to make improvement on batch material preparation and on information propagation part so that the cost of the preparation of delicious foods can be reduced and the value of the information can be enhanced. In this study, data mining technique is used to analyze the data collected in the laboratory so as to investigate whether the results of lean production and service as adopted by an enterprise can enhance the entire production and service performance; meanwhile, the characteristics of employees in the laboratory are investigated, and finally, lean service performance (LSP) and Employee Characteristic Analysis (ECA) models are set up to be used by the enterprise as reference for enhancing production and service performance and human resource utilization performance.

Fig. 1 Taiwan's delicious foods as prepared by laboratory

The main architecture of this article is as follows: The first section will be an introduction of the research motive and objective of this article. The second section will be an introduction of the data mining technique and related literature adopted in this article. The third section will be an introduction of the sample data and empirical analysis used by this article. In the fourth section, research conclusion and suggestion will be proposed.

2 Research method

2.1 Kmedoid cluster

Kmedoid is a newer cluster, and the main difference between Kmeans and Kmedoid stands in calculating the cluster centers: the new cluster center is the nearest data point to the mean of the cluster points.

2.1.1 Kmedoid algorithm

Given the data set X, choose the number of clusters $l < c < N$. Initialize with random cluster centers chosen from the data set X. Repeat for $l = 1, 2, \ldots$ Step 1: Compute the distances

$$D_{ik}^2 = (X_k - V_i)^T (X_k - V_i), \quad l \le i \le c, \, l \le k \le N \quad (1)$$

Step 2: Select the points for a cluster with the minimal distances, they belong to that cluster.

Step 3: Calculate fake cluster centers

$$v_i^{(l)*} = \frac{\sum_{j=1}^{N_i} X_i}{N_i} \quad (2)$$

Step 4: Choose the nearest data point to be the cluster center

$$D_{ik}^{2*} = (X_k - V_i^*)^T (X_k - V_i^*) \quad (3)$$

and

$$x_i^* = \arg \min_i (D_{ik}^{2*}); \, v_i^{(l)} = x_i^* \quad (4)$$

Until

$$\prod_{k=1}^n \max |V^{(l)} - V^{l-1}| \neq 0 \quad (5)$$

End: Calculate the partition matrix.

2.2 Grey relational analysis

Professor Deng [5] had proposed grey theory, which is nowadays successfully applied in all kinds of fields. Grey theory is a new theory, which is specifically

suitable for forecast and analysis. It has a characteristic to fully analyze finite data and information so as to achieve the goal of predicting the future data. Grey theory is mainly to, under the uncertainty of system model and the incompleteness of information, perform relational analysis and model construction related to the system; meanwhile, through forecast and decision-making method, the system situation is investigated and understood.

Grey relational analysis (GRA), in grey theory, is a measurement method for analyzing the relation among discrete sequence data, which is currently widely applied in performance assessment field [6–8]. It performs quantitative comparison and analysis on the development situation in the grey system factors. It is a method for evaluating the relational level among factors based on similar or dissimilar level of the development trend between factors, that is, all the relations among related factors of the system are displayed one by one to be used as useful information and reliable basis for system decision making and forecast and control; such analysis model can clarify the grey relation among factors in the grey system, meanwhile, it is a measurement that gives quantification to the trend of change of the system development.

The steps in the algorithm of Grey relational analysis are as follows:

Step 1: Find, from the original decision making matrix D, the standard sequence A_0 and inspected sequence A_i. Standard sequence is a set formed by the ideal target values of all the influential factors. $A_0 = (x_{01}, x_{02}, …, x_{0j}, …, x_{0n})$ contains j terms, wherein $j = 1, 2, …, n$. In addition, the performance value of inspected sequence $A_i = (x_{i1}, x_{i2}, …, x_{ij}, …, x_{in})$, wherein $i = 1, 2, …, m$.

Step 2: Normalize the data of original decision making matrix D.

Step 3: Calculate Grey relational distance Δ_{0ij}, and Δ_{0ij} is to evaluate the difference between each normalized data and normalized reference data.

$$\Delta_{0ij} = \left| x_{0j}^* - x_{ij}^* \right| \tag{6}$$

Step 4: Calculate Grey relational coefficient γ_{0ij}.

$$\gamma_{0ij} = \frac{\Delta_{\min} + \zeta \Delta_{\max}}{\Delta_{0ij} + \zeta \Delta_{\max}} \tag{7}$$

Step 5: Calculate Grey relational grade Γ_{0i}, with a formula as in the following:

$$\Gamma_{0i} + \sum_{j=1}^{n} [\omega_j \times \gamma_{0ij}] \tag{8}$$

Step 6: Make Grey relational ordinal and follow grey relational grade to perform performance ranking.

2.3 Genetic Programming

Genetic Programming (GP) is an algorithm developed by professor Koza [9, 10] in Stanford university of USA by following the genetic algorithm of Holland [11]. Similar to the concept of genetic algorithm, it has chromosome, fitness function, reproduction, crossover and mutation operation mechanism, etc. However, what is different is Genetic Programming further replaces the gene (0 and 1) of chromosome by syntax tree. Therefore, each individual in the matrix represents a set of computer program, and these program codes, similar to gene, through natural selection in the evolution process, can generate optimal program codes.

Genetic Programming originates from genetic algorithm. Usually, genetic algorithm uses binary string of fixed length to represent chromosome, and Genetic Programming uses tree structure that is highly changeable in size and shape and structure to represent chromosome and to represent, respectively, different formula. It is as shown in Fig. 2.

Here, the plus and minus signs are internal nodes, and the rest terminal nodes are elemental sets (X, Y, and 5) defined from the problem; and the representation formula corresponding to syntax tree is $(X - 5) + Y$.

In addition, in the model construction aspect, genetic algorithm uses known structure to find out optimal model parameter, but in Genetic Programming, the model structure is constructed together with model parameters. The basic formation components of Genetic Programming include:

1. Terminal set: It is a set formed by values sent into Genetic Programming. It can be constant or variable. Since it is located at the leaf part at the branch terminal of tree structure, it is thus called terminal set.
2. Function set: It includes arithmetic operation, logical operation, or the self-defined function by the user.
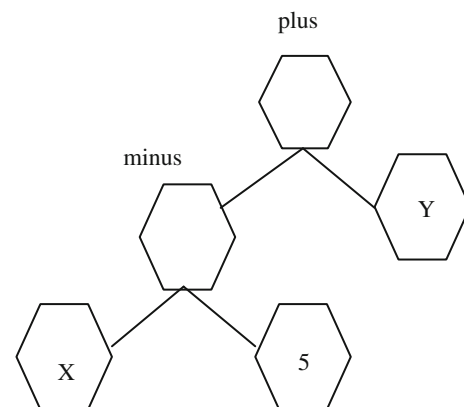

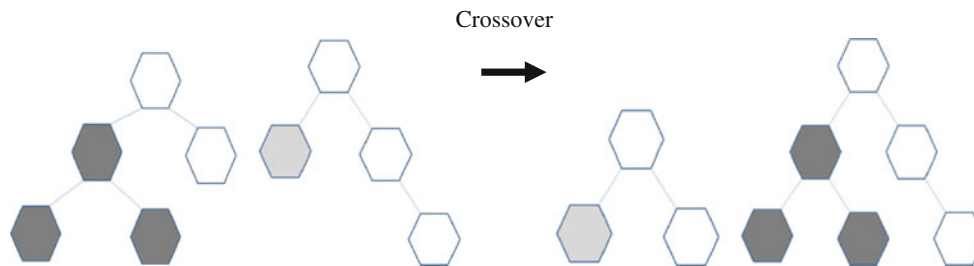Fig. 2 Syntax tree of Genetic Programming

Fig. 3 Crossover illustration

3. Fitness function: It is a formula designed to calculate and estimate the good or bad characteristic of chromosome. After fitness function operation is done on the chromosome, we will obtain a value. The closer the value to the target value, the higher the probability it is selected.
4. Control parameters: It includes population size, crossover probability and mutation probability, etc.
5. Termination condition: It can be set up as fixed number of generation or the error tolerance.

The main operator of Genetic Programming includes:

1. Selection: Based on the theory of natural selection of evolutionism of Darwin, before performing crossover and mutation, selection, and weed-out job will be performed first.
2. Reproduction: Excellent chromosome will be reproduced completely to the next generation from mother generation.
3. Crossover: To continue excellent gene in the chromosome of mother generation to the next generation. The operation method is to select in random two mother generations; meanwhile, nodes are selected in random as crossover points, then the parts below crossover points are exchanged to produce two different son generations. It is as shown in Fig. 3.

Mutation: It is to prevent the falling of chromosome into partial optimal solution, or to prevent the unlimited swelling of the layer number of chromosomes. In general operation, a node is selected in random as mutation point, next, operator will be selected in random from function set or terminal set to replace the original gene. But under certain conditions, under continuous increase of generation, chromosome will show a growth trend. In order to prevent tree layer expansion, mutation will start from random selection of gene in the tree, and the selected gene will be removed and replaced by gene right next to it, and all the parameters of the previous gene will be deleted. It is as shown in Fig. 4.

In Genetic Programming, after mechanisms such as selection, reproduction, crossover and mutation, the chromosome
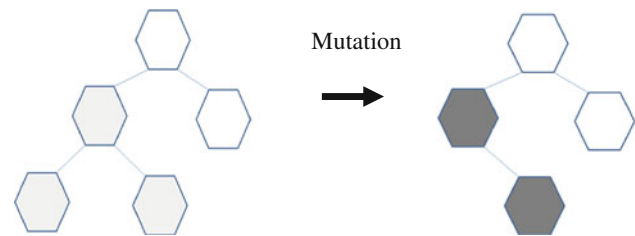

Fig. 4 Mutation illustration

in son generation will replace that in the mother generation, and this process will repeat again and again until the termination condition is met.

In this study, newer Matlab GPLAB toolbox will be adopted for model construction. During the model construction process, GPLAB will search dynamically related parameters of Genetic Programming (For example, crossover probability, mutation probability, tree depth, and node number) and use a chart to display the generation evolution process until the parameter is optimized. Currently, Genetic Programming has been applied in many fields, which include finance and banking field [12, 13], management field [14], and engineering field [15], and good results are obtained.

3 Empirical study

3.1 Sample data and variable

This study is carried out by Toyota lean service simulation laboratory at China University of Technology, and two sets of lean service experimental data are collected from March 23, 2010, to April 01, 2010. Here, in the first stage, there are 40 data, in the second stage, there are 40 data, and in the third stage, there are 20 data, that is, a total of 100 data. The related factors in each space include time spent from the arrival of the restaurant to the completion of dish ordering (X1), time spent from the completion of dish ordering to the arrival of ordered dishes (X2), time spent from the completion of dish ordering to the leaving of table for payment after eating (X3), time spent from the

Table 1 Descriptive statistical for the two sets of data

	Factors	X1 (sec)	X2 (sec)	X3 (sec)	X4 (sec)	X5	X6
First stage in combination with second stage (First set) $N = 80$	Max	300	450	630	735	3	2
	Min	3	16	90	210	0	0
	Avg	83.41	174.63	295.88	406.54	0.34	0.11
	Std	51.04	87.68	128.93	100.80	0.67	0.35
Second stage in combination with third stage (Second set) $N = 60$	Max	290	295	565	735	3	1
	Min	3	16	90	210	0	0
	Avg	79.30	142.33	213.97	379.72	0.22	0.83
	Std	60.87	66.92	91.00	88.03	0.55	0.28

completion of dish ordering to leaving after payment (X4), number of wrong dish sending (X5), and number of extra dish sending (X6). In this study, the research data in the first stage and second stage are combined to form the first set of data for a total of 80 data, then the research data in the second stage and third stage are combined to form the second set of data for a total of 60 data; this is to investigate the lean production and service performance and employee characteristic between two stages of the two sets of data, and the descriptive statistical values of these two sets data are as shown in Table 1.

3.2 The performance analysis of lean service simulation experiment

In this study, we referring to Pan [16] all the variables in two sets of data are used as indexes in lean production and service performance assessment, and the values of all indexes are the smaller the better. Then all the assessment indexes are analyzed using Grey relational analysis proposed by professor Deng [5], then the grey relational Matlab toolbox as developed by Wen et al. [17] is used to find out Grey Relational Grade for further sorting; the top 40 rankings in the first set and top 20 rankings in the second set are of good performance, and the latter 40 rankings are of bad performance, then it is compared with the previous stage (worse performance) and latter stage (better performance) within two sets so as to judge the performance of two sets of experimental data. The result of Grey relational analysis is as shown in Fig. 5. In this study, the minimal value in each space factor is used as standard sequence, wherein the bold black line at the bottommost side represents standard sequence, and the rest fine lines represent inspected sequence, that is, in each of the rest data, there are six nodes to represent the six space factors of that data. If inspected sequence gets closer to standard sequence, it means that the lean production and service performance of that data are better. After the analysis of this study, it was found that among latter 40 data of the first set of data, only five data are of bad performance; hence,
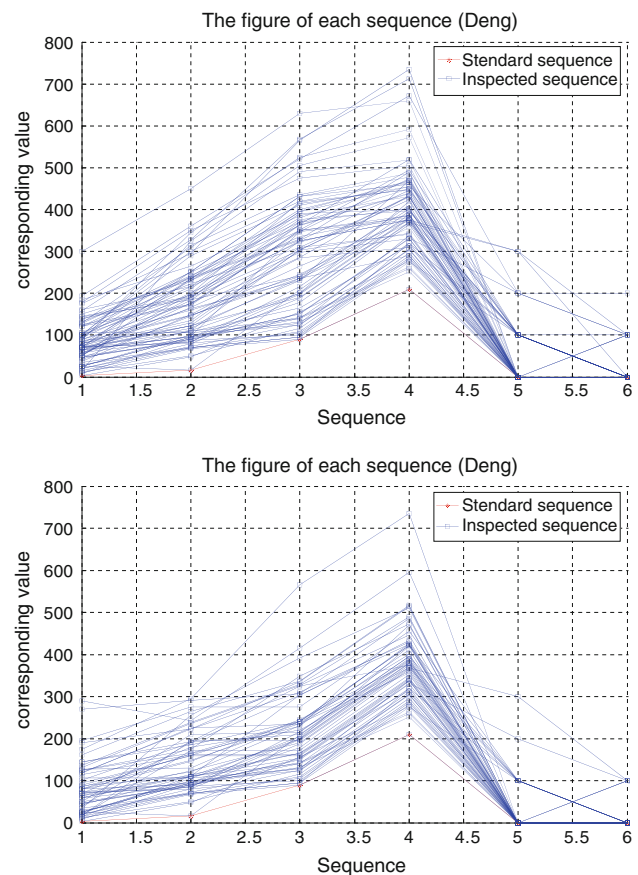

Fig. 5 The linear sequence chart of GRA

the performance in the second stage experiment is enhanced, as compared to that in the first stage experiment, by 87.5%; and in the latter 20 data of the second set of data, only six data are of bad performance; hence, the performance in the experiment in the third stage is enhanced by 70% as compared to that in the second stage, and from the analysis result, it can be judged that lean production and service can indeed enhance the entire service performance. Next in this study, this experimental data are used to set up lean service performance (LSP) model to be used as reference by the related enterprises; those enterprises only

need to substitute related factors into the model to judge whether it is necessary to take lean production and service to enhance the performance of that company. In this study, the latter 40 rankings of Grey relational analysis results of two sets of data represent bad performance (represented by 1), and the front 40 rankings in the first set and the front 20 rankings in the second set represent good performance (represented by 0); then this is used as dependent variable, and in association with six space factors, it is used as independent variable to perform the setup of three lean service performance model such as Genetic Programming, Back-Propagation Network (BPN) and logistic regression (LR).

First, the construction of Genetic Programming model is done first. In the model architecture, node mathematical function is formed by plus, mydivide, square, times, minus and mylog. The evolution initial parameter values in the generation include genetic generation number 100, 30 chromosomes, and the termination conditions are the number of generation. In this study, the first set and second set data are finely divided into 4 groups, wherein each small group data of the first set contains 20 data, and each small group data of the second set contains 15 data; meanwhile, three small groups are used as training sample data to construct the model, 1 small group is used as test data to test the model stability, and cross-verification is then carried out. Figure 6 is the output results using the front three small groups training data of the first set and second set data to construct Genetic Programming model, and Fig. 6a shows the dynamic adjustment tree depth and node number change in each evolution generation. Figure 6b shows Pareto front, which shows best fitness found for each tree size. Figure 6c shows Syntax Tree for Genetic Programming, which can be represented as the following two sets of Symbolic Regression.

The first set of regression formula is as follows:

plus(times(X3,plus(X3,times(X5,plus(plus(mydivide(mydivide(minus(X1,minus(plus(times(X5,plus(plus(X3,times(X6,X3)),X3)),plus(plus(X5,minus(X6,X3)),X5)),X3)),times(X5,plus(plus(X3,times(plus(X3,X3),times(X3,plus(X3,X5)))),plus(X3,times(X5,X3)))))),minus(plus(plus(X3,X4),X3),plus(X3,times(X5,X3)))),times(mydivide(minus(X1,minus(X6,times(X5,plus(X3,X5)))),X2),plus(X3,X3))),X1)))),times(mydivide(mylog(X4),mylog(X3)),plus(X3,times(X5,X5)))).

The second set of regression formula is as follows:

plus(X3,mydivide(X3,square(plus(X2,mydivide(plus(times(X4,mylog(X3)),mydivide(plus(X3,mydivide(X2,square(plus(X1,mydivide(plus(times(X4,mylog(X3)),mydivide(X2,minus(square(X2),X6))),mydivide(X4,X4)))))),X4)),X4))))).

Then in this study, Back-propagation Artificial Neural Network is further adopted to construct lean service performance model. In the architecture selection of neural network, this study has referred those mentioned in Yeh

[18], in the general issue, one hidden layer can be adopted, but for the neural number in the network hidden layer, it is feasible to adopt the division of input layer (6) plus output layer (1) by 2 and the round-off into 4, then we can obtain better forecast result. The internet architecture chart is as shown in Fig. 7a.

Finally, this study has adopted logistic regression in traditional statistical model to construct lean service performance model so as to compare three data mining techniques.

3.3 The Employee Characteristic Analysis of lean service simulation experiment

In this study, the X1, X2, X3, and X4 space factors within two sets of experimental data are used as Employee Characteristic Analysis (ECA) indexes, wherein X1 time spent from the arrival at the restaurant to the completion of dish ordering, which represents whether the attitude the employee serving the customer is aggressive enough; X2, X3, and X4 are time spent from the completion of dish ordering to the arrival of ordered dishes, time spent from the completion of dish ordering to the leaving of table for payment after eating, and time spent from the completion of dish ordering to leaving after payment, and all the factors within these spaces are related to whether the employees are familiar with the preparation of dishes or whether the dishes serving is of good production efficiency. In this study, these two sets of data will be done with Kmeans and Kmedoid hard clustering analysis, and the analysis result will divide the data into two Clusters; meanwhile, the writing of professor Liu and Xu [19] will be referred to and newer Fuzzy Sammon Mapping non-linear mapping will be adopted to map the clustering result from multidimensional space to two-dimensional plane so as to clearly display the clustering result, which is as shown in Fig. 8.

From a comparison of two sets of data, it was found that one group has better service attitude and has dish-serving efficiency, and another group has worse service attitude and is lacking of dish-serving attitude; however, from the clustering index, it was found that in the first set, the Kmeans PI (Partition Index) is 1.7732, SI (Separation Index) is 0.0222, which is much higher than Kmedoid PI index of 0.0025 and SI index is 0.000003; in the second set, Kmeans PI index is 1.0448, SI index is 0.01742, which is much higher than Kmedoid PI index of 0.0029 and SI index of 0.000005; since PI and SI index is the smaller the better, further from the figure, we can find that Kmedoid clustering result can clearly separate two groups, and there is obvious gap between groups; hence, in this study, Kmedoid clustering result is used for subsequent analysis. It was found from the first set of clustering result that there are 50

Fig. 6 The output charts of
LSP model from GP
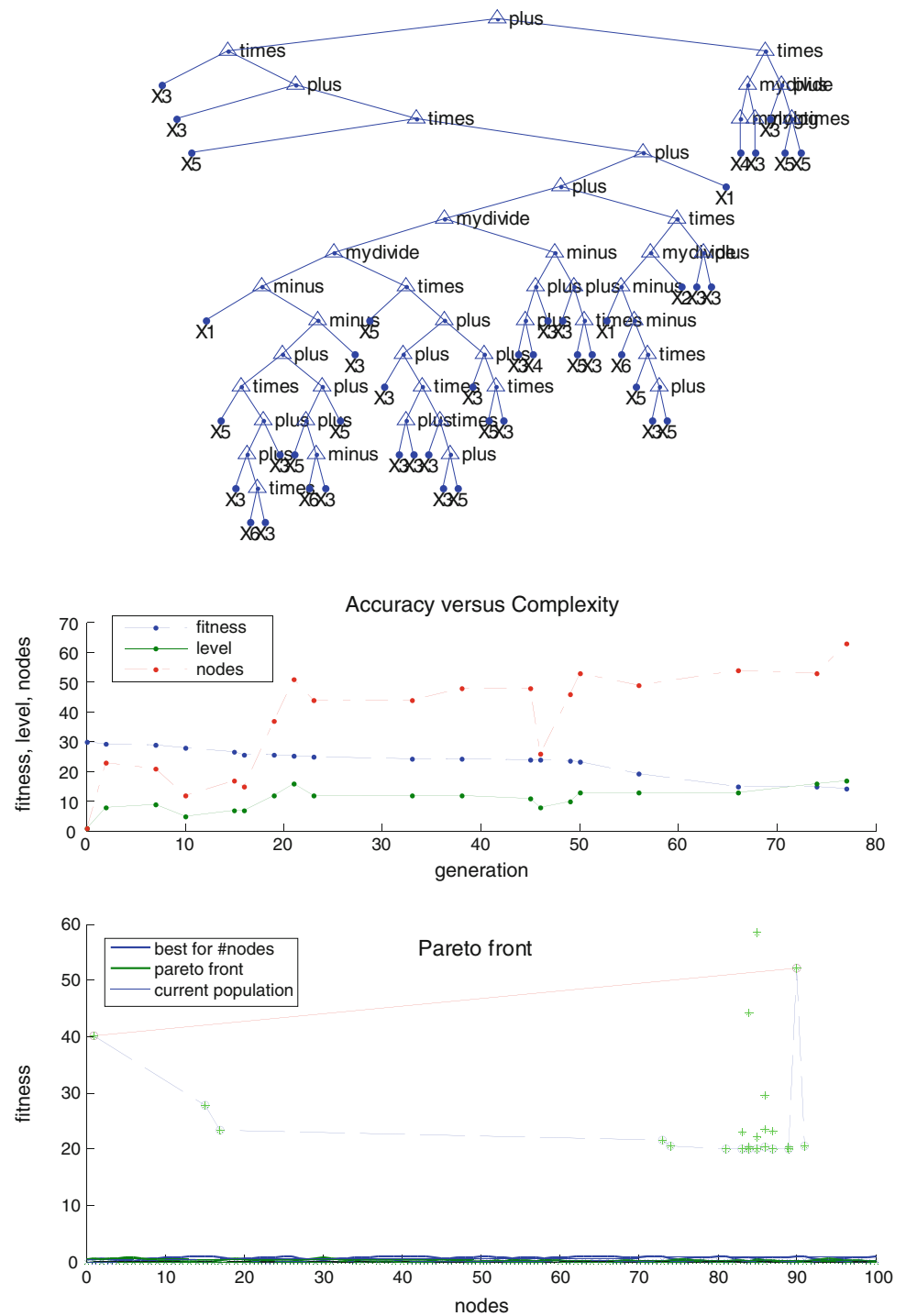
data classified as better service attitude with dish-serving efficiency (represented by 0), and there are 30 data classified as bad service attitude and bad dish-serving efficiency (represented by 1). In the second set of clustering result, 42 data are classified as better service attitude and good dish-serving efficiency (represented by 0), and 18 data are classified as bad service attitude and worse dish-serving efficiency (represented by 1).

Next, in this study, the previous clustering result is used as dependent variable, X1, X2, X3, and X4 space factors are used as independent variables to perform the construction of three Employee Characteristic Analysis models such as Genetic Programming, Back-Propagation Network and logistic regression so as to be used as reference by enterprises, and the enterprise can then detect the production and service performance of the

Fig. 6 continued

Pareto front

Accuracy versus Complexity
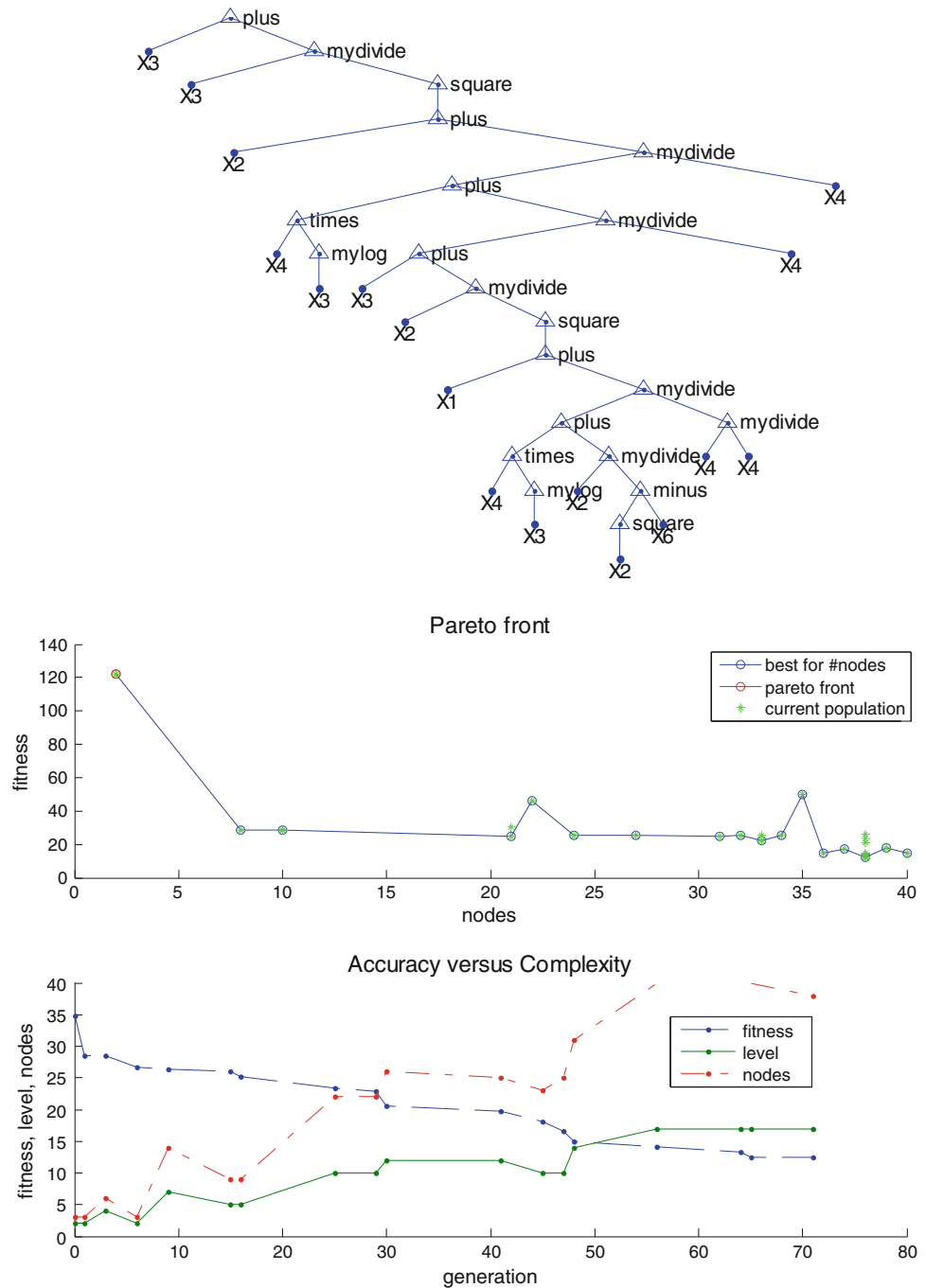
employee at any time. The same as the above section, data in the first set and second set are finely divided into four groups, wherein there are 20 data in small group in the first set, and 15 data in small group in the second set; meanwhile, cross verification is carried out in a way that three small groups are used as training sample data to construct a model and one small group is used test data to test the model stability. The selection of model parameter of Genetic Programming is the same as the last section. Figure 9 shows the output result using the first three small groups training data of first group and

second group data to construct Genetic Programming model, from Fig. 9a, it can be seen the dynamic adjustment of tree depth and node number change in each evolution generation. Figure 9b shows Pareto Front, which is the best fitness found for each tree size. Figure 9c shows the Syntax Tree of Genetic Programming, which can be represented as the following two sets of symbolic regressions.

The first set of regression formula is as follows:

mylog(minus(X4,times(plus(plus(minus(square(times(X6, plus(X4,X4))),X4),times(square(mylog(minus(X4,times(plus

Fig. 7 The network structure charts of LSP model from BPN
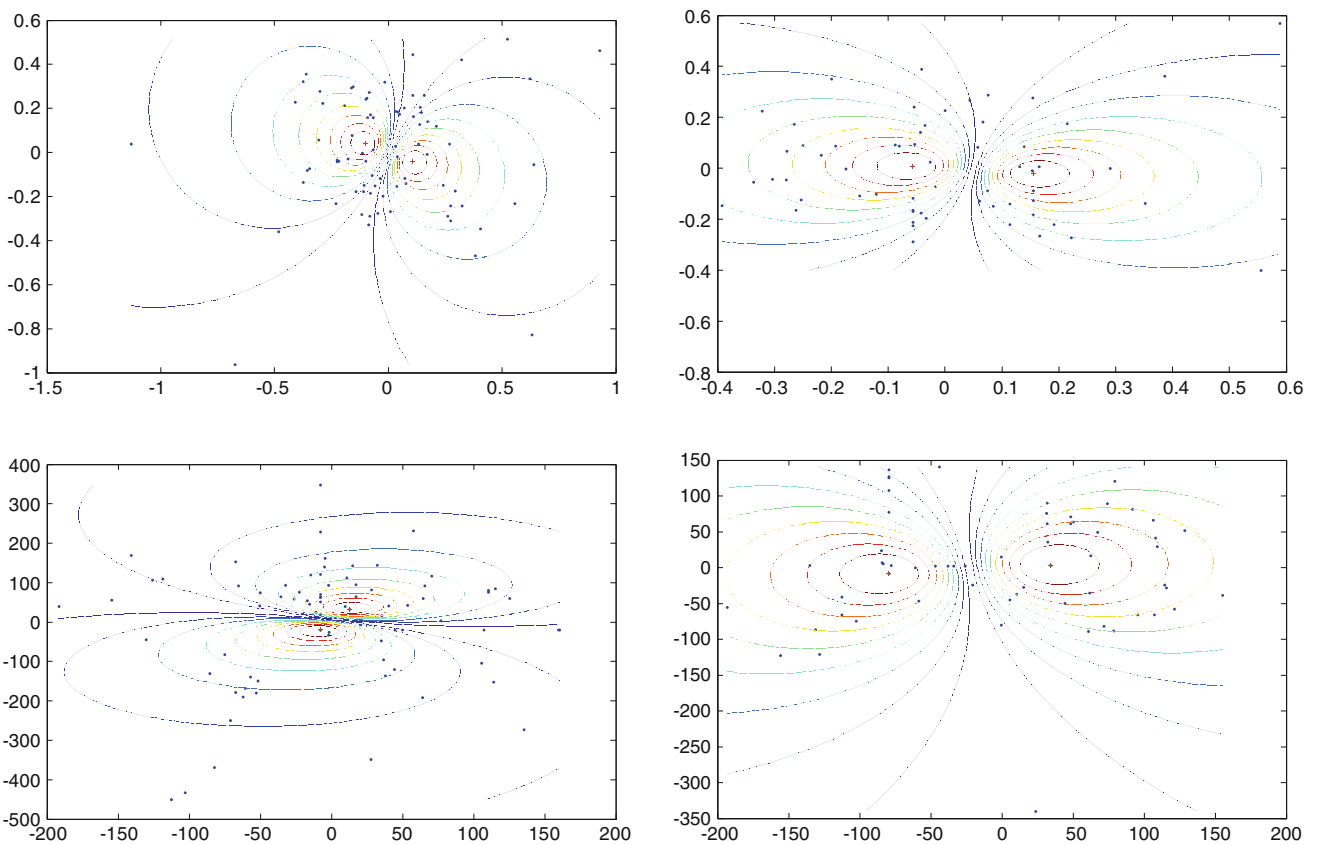
Fig. 8 Kmeans and Kmedoid clustering analysis result
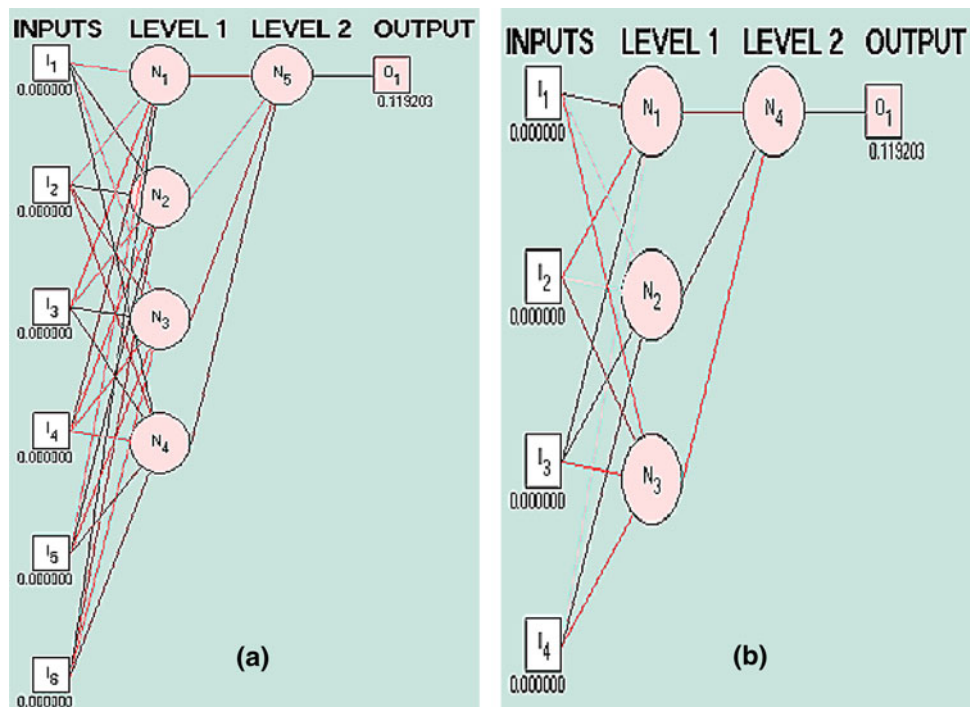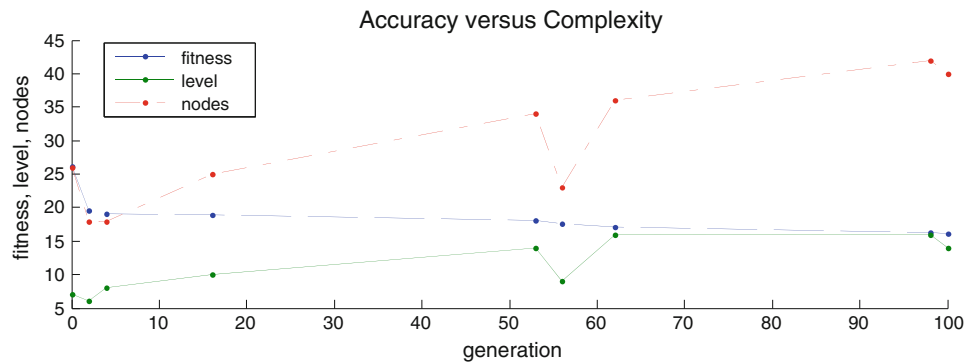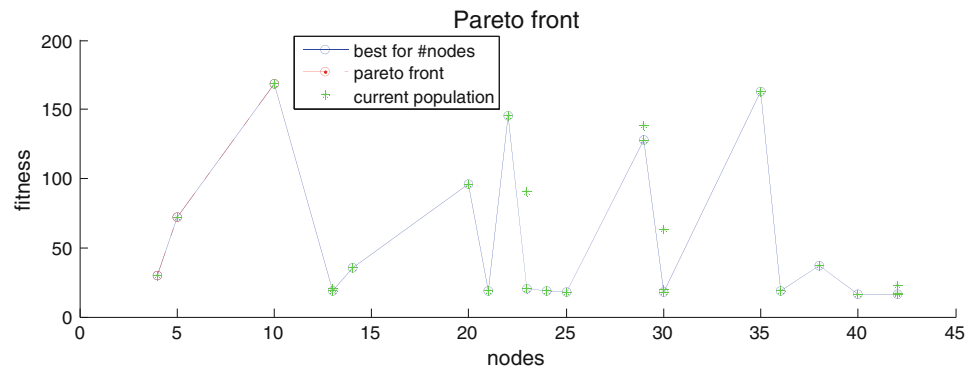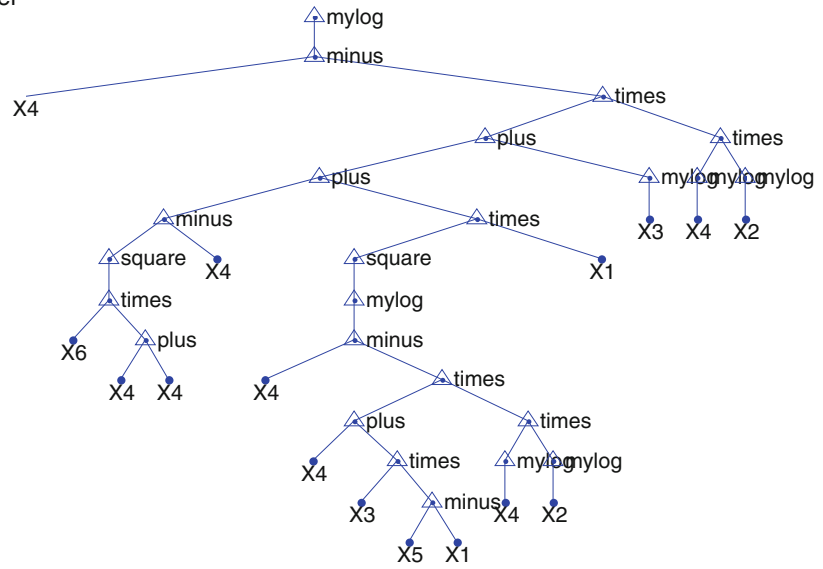
Fig. 9 The output charts of ECA model from GP

(X4,times(X3,minus(X5,X1))),times(mylog(X4),mylog(X2)))))),X1)),mylog(X3)),times(mylog(X4),mylog(X2))))).

The second set of regression formula is as follows:

minus(times(X2,mylog(times(mylog(plus(X2,minus(plus(X6,mylog(mylog(times(plus(X4,plus(X6,X2)),plus(X6,plus(X6,X2)))))),square(plus(X2,X4))))),mylog(mydivide(X4,mylog(X5)))))),minus(X2,square(square(X6)))).

Then in this study, Back-propagation Artificial Neural Network is further adopted to construct Employee Characteristic

Analysis model. In the architecture selection of neural network, one hidden layer is adopted in this section; but for the neural number in the network hidden layer, it is feasible to adopt the division of input layer (4) plus output layer (1) by 2 and the round-off into 3. And the network architecture is as shown in Fig. 7b. Finally, in this study, the logistic regression of traditional statistical model is adopted to construct Employee Characteristic Analysis model so as to compare the advantages and disadvantages of three data mining techniques.

Fig. 9 continued

Pareto front

Accuracy versus Complexity

3.4 Summarization and comparison of the classification and forecast capabilities of three models

In this study, the output result of Genetic Programming model and Back-propagation Artificial Neural Network model is defined that when the value is smaller than 0.5, it is classified as 0, when the value is larger than 0.5, it is classified as 1; and this is to observe the forecast capabilities of three models. After cross verification of the four small groups of data with the results generated by three models, a ROC curve is then plotted, which is as shown in Fig. 10.

Figure 10 shows the ROC curve of classification forecast result of cross verification of four small groups of data of two sets of data in lean service performance model and Employee Characteristic Analysis model. Bradley [20] pointed out that the larger the area above reference line and underneath the curve, the more accurate the classification capability of the model. From the figure, it can be

Fig. 10 The ROC curve of classification forecast results for LSP and ECA model

obviously seen that no matter it is lean service performance model or Employee Characteristic Analysis model, Genetic Programming model always shows the best classification capability.

Then from Table 2, the output result of ROC curve analysis can be seen. Here, sensitivity (Sen) means the percentage of the ratio of the number of forecast result of 1 to the number of real value of 1; and specificity (Spe) means the percentage of the ratio of the number of forecast result of 0 to the number of real value of 0. Moreover, Hand and Till [21] had pointed out that (Gini Index) $= 2 \times$ AUC $- 1$. Since the larger these index values the better, for the specificity and sensitivity of Genetic Programming model in the table, area under the curve (AUC) and Gini Index, no matter in lean service performance model or Employee Characteristic Analysis model, are all higher than that in other model. Hence, it has very good alarming and detection capabilities.

4 Conclusions and suggestions

The main contribution of this study is to further investigate on whether the experimental data obtained from Toyota lean service simulation experiment at China University of Technology have production and service performance. From Grey relational analysis, it was found that the result after lean production process (the second stage) can indeed be the entire production and service performance. In addition, this article also investigates employee characteristics during the production process. From the clustering analysis, it was found that employee characteristic can be divided into two groups of good service attitude with dish-serving efficiency and of worse service attitude and without dish-serving efficiency. Finally, this research has further set up lean service performance model and Employee Characteristic Analysis model to be used by the enterprise as reference for enhancing production and service performance and human resource utilization performance.

In addition, this article uses Grey relational analysis to perform the performance analysis of lean production and service. In the future, it is suggested that other models (for example, Data Envelopment Analysis) can be

Table 2 The empirical result of ROC curve analysis

	Model	Set	Sen	Spe	Auc	Gini	Set	Sen	Spe	Auc	Gini
LSP	GP	1	0.950	0.925	0.938	0.876	2	0.950	0.950	0.950	0.900
	BPN		0.900	0.825	0.863	0.726		0.900	0.800	0.850	0.700
	LR		0.825	0.775	0.800	0.600		0.850	0.800	0.825	0.650
ECA	GP	1	0.980	0.967	0.973	0.946	2	0.944	0.952	0.948	0.896
	BPN		0.860	0.867	0.863	0.726		0.889	0.805	0.897	0.794
	LR		0.900	0.833	0.867	0.734		0.833	0.833	0.833	0.666

adopted to further investigate the production and service performance.

References

1. Womack JP, Jones DT (2006) Lean thinking. Simon & Schuster, New York
2. Arnheiter E, Maleyeff J (2005) The integration of lean management and six sigma. TQM Mag 17(1):5–18
3. Su CT, Chiang TL, Chang CM (2006) Improving service quality by capitalizing on an integrated lean six sigma methodology. Int J Six Sigma Compet Advant 2(1):1–22
4. Chen JC, Lee H, Chiu S, Tseng B (2011) Productivity improvement with lean production in glove manufacturing industry. Key Eng Mater 450:247–250
5. Deng J (1982) The control problems of grey system. Syst Control Lett 5:288–294
6. Kuo YH, Huang ST (2007) The association of grey relational analysis and data envelopment analysis method to evaluate the enterprise performance of steel industry, 43rd annual meeting and 13th nationwide quality management forum of Chinese Society for Quality
7. Yen JH, Wang SA (2005) Evaluation of operational performance for domestic financial holding companies using the entropy theory and grey relational analysis, 2005 third "Management thinking and practice" academic forum
8. Lin SY (2004) Evaluation of business reputation in information service industry—an application of grey relational analysis. J Inf Technol Soc 2:79–95
9. Koza JR (1992) Genetic programming: on the programming of computers by means of natural selection. MIT Press, Cambridge
10. Koza JR (1992) Genetic programming II: automatic discovery of reusable programs. MIT Press, Cambridge
11. Holland J (1975) Adaptation in natural and artificial systems. University of Michigan Press, Ann Arbor, MI
12. Noe TH, Wang J (1997) The self-evolving logic of financial claim prices. Genetic algorithms genetic programming computational finance, pp 249–262
13. Kaboudan M (2000), Using GP forecasts to enhance profitable trading of stocks. In: Proceedings of the 5th joint conference on information science, pp 925–928
14. Chen SH, Chie BT (2006) Automatic management of innovation: the application of evolutionary computation method. J Technol Manag 11(2):97–126
15. Jan JF (1998) Genetic programming for classification of remote sensing data. Taiwan J For Sci 13(2):109–118
16. Pan WT (2010) The use of genetic programming for the construction of a financial management model in an enterprise. Int J Appl Intell, Published Online: 23 Oct 2010
17. Wen KL, Chang-Chien SK, Yeh CK, Wang CW, Lin HS (2006) Apply MATLAB in grey system theory. Chuan Hwa Book CO., LTD, Taipei
18. Yeh YC (2001) The model application and practice of artificial neural network. Scholars Books Co., Ltd, Taipei
19. Liu L, Xu W (2006) UOFC-AINet: a fuzzy immune network for unsupervised optimal clustering, cimca, pp. 196, International conference on computational intelligence for modelling control and automation and international conference on intelligent agents web technologies and international commerce (CIMCA'06)
20. Bradley AP (1997) The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern Recognit 30(7):1145–1159
21. Hand DJ, Till RJ (2001) A simple generalisation of the area under the ROC curve to multiple class classification problems. Mach Learn 45(2):171–186