

# Predicting Credit Card Defaults with Deep Learning and Other Machine Learning Models

Tsungnan Chou and Mingmin Lo

**Abstract**—Since the overdue amount of credit cards has been increasing year by year, the rising credit card delinquencies might prevent the commercial banks to allocate more funds in profitable investments. At the same time, the high processing costs of credit card delinquencies through manual verification also affect the competitiveness of credit card issuers. Because the market competition becomes strict in the era of financial technology, to predict correctly whether cardholders will be unable to pay off credit card debt and establish an effective risk prediction model is the major purpose of this study.

We first implemented four machine-learning approaches to predict the default cases, however, most models encountered challenges to resolve imbalance problem of delinquency cases in data sets and reported lower predictive accuracy. Two inference strategies including grey incidence analysis and fuzzy decision tree were proposed to improve the predictive performance. The average accuracy for both strategies were increased from 0.82 to 0.86 and 0.89 respectively. In addition, the deep learning approach integrated with various network structures was also incorporated to evaluate model performance. The experiment results indicated the deep neural network performed better in most evaluation metrics and achieved an impressively high accuracy of 0.93 if compare to the machine learning models. Finally, three feature selection methods were employed with the deep learning model, and the results showed similar predictive accuracy as the original deep learning models with slightly better performance being reported by filtering variables with the grey incidence analysis. This research work could be extended to apply more complicated deep learning algorithms to learn and trace the behaviors of the credit card holders and reduce the default risks for banking industries.

**Index Terms**—Default prediction, machine learning, deep neural network, deep learning.

## I. INTRODUCTION

As more and more consumers rely on the credit card to pay their everyday purchases in online and physical retail store, the amount of issued credit cards and the overwhelming amount of credit card debt by the cardholders have rapidly increased. Therefore, most financial institutions have to deal with the issues of credit card default in addition to the credit card fraud such as credit card dump. Both the credit card verification applied to the cardholders and the default risk management after card issued are crucial to the future success of most financial institutions. Under those circumstances, to reduce the delinquent payment of credit card holders, and perform the effective management of credit card risk will help financial and banking institutions to achieve competence

advantages.

A default analysis system empowered with data mining techniques will allow the credit card issuers to monitor and eliminate the ongoing transactions of potential risky account prior to credit fraud and default. Many researchers have developed various risk prediction systems to detect the credit card defaults for banking industry, some research works such as [1], [2] created the default prediction systems based on statistic techniques. More studies that are recent have reported that Artificial Intelligence (AI) methods might perform better than the statistical methods for credit risk assessment. Data mining techniques and machine learning algorithms can assist data managers to improve their problem-solving techniques by discovering meaningful patterns and trends from large amounts of data. Meanwhile, the fast development of artificial intelligence and deep learning techniques also bring great impacts to the financial industries.

This study aims to compare the prediction accuracy for the newer developed deep learning methods against with the conventional machine learning based on the financial information of the credit card holders.

The implementation of the framework comprised three stages. The first stage applied conventional machine learning techniques including decision tree (DT), naive Bayes (NB), K nearest neighbor (KNN), and logistic regression (LR) to implement default inference systems for predictive tasks of financial institutions. A modified fuzzy decision tree (FDT) was also implemented with different number of membership functions to evaluate whether the fuzzy partition and K-Means clustering could improve the predictive performance. The experiment result from our previous study [3] applying grey incidence analysis (GIN) was included in this paper for comparison with the KNN models. The second stage used a deep learning system based on the customized modification of layers, neurons and activation functions. The last stage combined three feature selection methods to reduce the dimensionality of input variables for the deep neural models.

The overall objective of this framework intended to extract useful features from cardholder data, and applied the artificial intelligence techniques to assist financial institutions to reduce their risk of credit card default. Both the personal information and repayment details of the cardholder were used to trace the consumer behavior for the predictive models. The rest of this paper was organized as follows. Section II documented the proposed models and related methodologies. Section III discussed the data processing and evaluated experimental results, and finally the Section IV provided a conclusion of the work.

Manuscript received July 1, 2018; revised August 17, 2018.

Tsungnan Chou and Mingmin Lo are with the Dept. of Finance, Chaoyang University of Technology, Taichung County, 41349 Taiwan, R.O.C. (e-mail: tnchou@cyut.edu.tw, mingminlo@cyut.edu.tw).

## II. RESEARCH AND EXPERIMENT DESIGN

For examining the accuracy of the predictive models, the datasets applied in our previous research work [3] were imported as the training and testing data. The datasets consisted of the credit card holder's personal and transaction information with a mix of categorical and continuous attributes, and the dependent attribute gave a value of zero if there was no default and a value of one if there was a default. In this study, there were totally 24 variables expressed as the account and repayment information of the holders. The raw data were segmented into default records and non-default records and then were randomized to avoid sampling bias. To provide sufficient and adequate data for the evaluation of models, 2000 credit card accounts were reviewed and subdivided into 1000 for the training dataset and another 1000 for the testing dataset. Considering the data with differed units and scale, the training and testing datasets were normalized or standardized to a common scale according to the property of variables prior to the training process.

The methodologies established to predict the default of credit card holders included the machine learning (ML) and deep learning (DL) approaches, and each approach was examined with several individual models. The first experiment employed four machine-learning methods and worked as the benchmark models, because their results were compared with other modified models to examine whether our strategies would improve the predictive performance. To compare performance with the KNN model, the experiment result in our previous work [3], which used the grey incidence analysis as an instance-based learning approach, was addressed in the second experiment. Since the decision tree applied in the first experiment failed to compete with the performance of other models, the fuzzy decision tree was involved in the third experiment to examine whether the combination of fuzzy theory and decision tree improved the predictive performance. In addition to the conventional ML models, the newly developed deep learning approach was evaluated with different number of layers and neurons in another two experiments.

As the original ML models used all 24 variables as inputs, high-dimensional data might increase the analytical difficulty for machine-learning models. Therefore, three feature selection methods were applied in the last experiment to reduce the number of input variables for models. Especially, the grey incidence analysis was also incorporated as one of the feature selection due to its ability of ranking the relational order of variables. The methods and strategies used in this study were briefly described as follows.

### A. Machine-Learning Models

Four most common used machine-learning techniques including logistic regression, naive Bayes, decision tree, and K-nearest neighbor methods addressed in [4] were employed to predict the delinquency cases of credit card based on the original 24 variables. The logistic regression was a statistics methodology used to account for the relationship between variables; especially the dependent variable was categorical or binary. On the other hand, naive Bayes method was a supervised learning algorithm based on Bayes theorem and worked very well in many research applications. The

approach only required a small amount of training data to estimate the required parameters. The C4.5 decision tree [5], improved from the original ID3 algorithm could handle training data with missing values and performed classification tasks with improved computational efficiency.

Moreover, K-nearest neighbor was another supervised learning technique mostly used for classification tasks such as the default and non-default cases. As the KNN searched the entire training dataset for the K most similar cases from the neighboring instances and provided a summarizing outcomes for those K neighbors, this method was vulnerable to erroneous and outlier data. In comparison to the KNN method, the grey incidence analysis measured the closeness of a relationship between sequences by evaluating the similarity of the geometrical curves formed by the sequences, which were considered as an analogous to the corresponding K neighbors in KNN model using different distance measures. The prediction of a testing sample by KNN model was fulfilled with the five nearest neighbors based on Euclidean distance in this study. Meanwhile, the GIN model calculated the metric of incidence degree for each variable in sequences, and then used the summarized metric to represents the geometrical similarity between all variables. As a result, the more similar for the geometrical shape of curves formed by all variables, the higher grey incidence degree between them. For the purpose of providing comparison with different searching algorithm and feature selection, the predictive result of the GIN model in our previous experiment [3] was involved as a counterpart for performance comparison with KNN model. The GIA-A model used the original 24 variables. In addition, the GIA-B model comprised 10 out of 24 variables selected by four different incidence measurements fusing with evidential reasoning approach. Moreover, GIA-C model derived from the GIA-B model but worked with a dynamic weighting.

In addition, for the purpose of improving the performance of decision tree in the first experiment, another fuzzy decision tree model based on fuzzy partition and K-Means clustering [6], [7] was also employed as a counterpart for comparison to the classical decision trees. Both the input and output variables were partitioned by various fuzzy membership functions. The maximum partition size and the number of membership was defined by K-Means clustering method for each active variable. The K-Means method was an unsupervised learning technique used for clustering, and K was the number of clusters the method tried to learn from the dataset. The fuzzy decision tree was also pruned to reduce its complexity, and the paths from the root node to the leaf nodes were converted as decision rule [8]-[10]. The FDT model was examined with three partition methods, which were denoted as FDT-M3, FDT-M5, and FDT-M7 models respectively.

### B. Deep Learning Models

The conventional artificial neural network built with multiple hidden layers has been successfully solving various image classification and recognition problems under complex scenarios. Each hidden layer could learn the abstraction of previous layer output. Especially, the deep learning algorithms and models have been well developed for many years, several important networks, such as deep neural

networks (DNN), convolutional neuron networks (CNN), recurrent neural networks (RNN), autoencoder (AE), were pervasively applied all over the science and technology [11], [12]. An autoencoder neural networks, including compression and decompression functions, could be stacked to learn the representation of variables without assigning labels for supervised learning. A recurrent neural network was normally employed with the application based on time sequence data. On the other hand, the convolutional neuron networks normally was used to learn patterns and components of an image. However, training neural networks with multiple hidden layers could be time-consuming and difficult if the structure of network became more complex. The deep neural networks applied in this study included several steps. The first step was to initialize a sequential DNN model consisted of different stack of layers, and followed by appending customized input layer, hidden layers and an output layer to the model according to what problems the model tried to solve in application.

As the hidden and output layers controlled the DNN inner workings, the number of hidden layers and their corresponding processing units might affect the performance of predictive model. The hidden layers within the sequential model were experimented with 2 to 5 layers, and the processing units were tested for 64, 32, 16 units respectively. Furthermore, we also used dropout layers and set up cutoff threshold value to control over-fitting problem in this study. The dropout rate was set to 0.2 to remove the processing units with weights below 20%. Both the Sigmoid and ReLU activation functions were used in different layers, and the max-pooling layer, where a small 2\*2 pooling region acted as a filter was applied to filter the output of the previous layer and reduced the dimensionality of input pattern by eliminating non-maximal values. Finally, the last layer was a fully connected layer and similar to a traditional neural network. The Softmax function employed in this layer classified the outcomes into the final default categories. The implementation of deep neural network was fulfilled by the TensorFlow open-source software library and Keras API based on the following processes.

Step 1: Select the top N variables as inputs based on feature selection.

Step 2: Construct a sequential deep neural network and initialize the sequential model.

Step 3: Define the number of hidden layers and corresponding network nodes.

Step 4: Add dropout layers and specify the dropout rate.

Step 5: Assign the ReLU and Sigmoid activation functions for layers.

Step 6: Reshape the output variable with one-hot encoding.

Step 7: Apply a fully connected output layer.

### C. Feature Selection Methods

To evaluate the influence of feature selection on accuracy of deep learning models, the grey incidence analysis, Pearson correlation analysis, and information gain were applied in the final experiment. By organizing the 24 input variables as data series, the grey incidence analysis measured the similarity between the testing and training data, and identified the most similar pattern matched within the training data series by

calculating incidence coefficient for each variable in the two observed data series [13]. The individual coefficients were aggregated as the degree of grey incidence to represent the overall measurement of two data series according to different methodologies described in our previous study. On the contrary, a data series could also be formed by all instances collected by a specific variable, and the degree of incidence was then used to rank the relationship between variables. If we used grey incidence analysis to compare series of instances rather than that of variables, then we could acquire a ranking list of variables based on the calculated degree of incidence. The Pearson correlation analysis explored the strength and direction of linear relationships between the variables and was another alternative for feature selection. Despite this method required the assumption that the variables were normally distributed and encountered with the problem of multicollinearity to select relevant variables, the result of correlation test could be used in our experiment to choose the top ten variables as inputs, which were highly relevant to the dependent variable for models. Additionally, information gain was another simple option to get the important attributes for feature selection. The entropy of each input variable was calculated for the dependent variable, and the input variables that contribute more information were selected.

## III. EXPERIMENT RESULTS

The experiment results were divided into three parts in this section. The first part comprised three experiments based on conventional machine-learning models, but incorporated different modifications to improve predictive performance. Moreover, the second part implemented the deep learning technique with various network structure and contained another two experiments. The third part employed three different feature selection methods with the deep neural network models. Totally four performance metrics derived from the confusion matrix were used to evaluate the model performance including accuracy, sensitivity, specificity, precision, and F-score.

### A. Performance of the ML Models

As shown in Table I, four conventional machine-learning models were evaluated for the prediction performance, and among all models, the DT model was inferior to other three models in accuracy metric although all model achieved the accuracy over 0.76. Because 76% of the collected default cases in dataset were positive, therefore, we dealt with the imbalanced dataset in this study. As the result, the DT model obtained the lowest accuracy and was not regarded as better than the accuracy of random guessing.

TABLE I: RESULTS OF CONVENTIONAL ML MODELS

| Model | Accuracy | Precision | Recall | F-Score |
|-------|----------|-----------|--------|---------|
| DT    | 0.762    | 0.762     | 1.00   | 0.865   |
| KNN   | 0.806    | 0.843     | 0.916  | 0.878   |
| NB    | 0.909    | 0.998     | 0.894  | 0.943   |
| LR    | 0.785    | 0.840     | 0.886  | 0.862   |

To demonstrate the performance of search algorithm for neighboring instances against the KNN model, the

experiment result presented in our previous study using the grey incidence analysis was summarized in Table II. The results showed all of the performance metrics of GIA models obtaining better outcomes than most models in the first experiments, despite the fact that the NB model reached a moderately higher accuracy at 0.91. Meanwhile, both GIA-B and GIA-C models achieved higher accuracy and were considered superior to GIA-A model in exhibiting fewer variables. In contrast to GIA-C, GIA-B model achieved a higher specificity, but the GIA-C model acquired higher sensitivity. In banking industries, if considering the default cases were the critical issue resulting the serious loss to banking in terms of false negatives, then sensitivity was a very important metric and the GIA-C model might be better choice than other models.

TABLE II: RESULTS OF GIA MODELS

| Model | Accuracy | Precision | Recall | F-Score |
|-------|----------|-----------|--------|---------|
| GIA-A | 0.838    | 0.691     | 0.917  | 0.635   |
| GIA-B | 0.856    | 0.698     | 0.904  | 0.701   |
| GIA-C | 0.891    | 0.946     | 0.989  | 0.718   |

Since the predictive accuracy of DT model was considerably worsened than other machine-learning models in the first experiment, the FDT models based on fuzzy partition and K-Means clustering were involved as another alternative to improve the predictive accuracy. As opposed to the original DT model, all FDT models presented similar improvement and increased the average accuracy to 0.887 in Table III, a 12% difference from the original DT model being reported. The FDT-M7 model reached a slightly lower accuracy at 0.882.

TABLE III: RESULTS OF FDT MODELS

| Model  | Accuracy | Precision | Recall | F-Score |
|--------|----------|-----------|--------|---------|
| FDT-M3 | 0.890    | 0.874     | 0.998  | 0.932   |
| FDT-M5 | 0.893    | 0.887     | 0.984  | 0.933   |
| FDT-M7 | 0.882    | 0.893     | 0.959  | 0.925   |

### B. Performance of the DL Models

In comparison to the performance of machine learning models, deep neural network models were evaluated with different structures in experiments. The DNN models were implemented with two, three, four and five hidden layers respectively, and each layer comprising the number of neurons from 64 to 16. The training process of deep learning model assigned 50 for the batch size, which indicated the number of samples used to update the training gradient within each epoch. The total number of training cycles was set to 40 epochs. Fig. 1 shows the loss and accuracy curves in the top and bottom panels during training process. The training loss decreased and became smooth after 15 epochs, while the training accuracy increased and became smooth after the similar epochs. Both the training and validation loss curves decreased together without divergence between them. On the other hand, the accuracy curve for both began to increase after few epochs and then reached the highest over 90%. The Fig. 1 also indicated the use of dropout layer was capable of preventing overfitting while the dropout rate was set to 0.2.

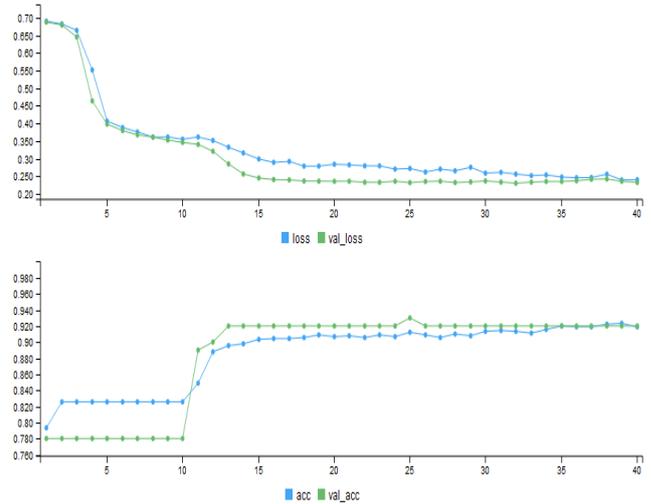


Fig. 1. The loss and accuracy curves during training process.

To examine whether the number of hidden layers would increase the predictive accuracy for deep learning models, four different number of hidden layers were integrated with DNN models, including two, three, four and five layers respectively. All 24 variables were selected as inputs to enable provably consistent configuration with the ML models. The fourth experiment evaluated the performance of deep learning model with various stack layers. By referring to Table IV, the precision metric of DNN-3L was better than other models. If we compared models in accuracy metric, all models performed well, and especially the DNN-3L model reached the highest accuracy of 0.932. However, the recall metric dropped slightly by 0.025 if compared to DNN-5L model which was 0.754. All models in this experiment achieved the substantial high accuracy greater than 0.920.

TABLE IV: RESULTS OF CONVENTIONAL DNN MODELS (VARIOUS LAYERS)

| Model  | Accuracy | Precision | Recall | F-Score |
|--------|----------|-----------|--------|---------|
| DNN-5L | 0.930    | 0.943     | 0.754  | 0.838   |
| DNN-4L | 0.926    | 0.932     | 0.746  | 0.829   |
| DNN-3L | 0.932    | 0.983     | 0.729  | 0.837   |
| DNN-2L | 0.931    | 0.972     | 0.733  | 0.836   |

The hidden layer was constructed with the minimum layers in the sixth experiment to examine the model performance affected by the number of neurons, which were assigned to 64, 32, and 16 respectively. The experiment results in Table V showed that the best result was DNN-64N model with the accuracy of 0.931, which outperformed other models. However, this result implicated that the number of neurons in each layer was not sensible for 64 and 32, and the increase of the hidden layers and neurons might not be able to improve the predictive accuracy substantially.

By referring to Table V, the recall metric of DNN-64N model was better than the other two models. The result also indicated that the default cases were correctly classified by DNN-64N model with a probability of 73.3%. This model used more neurons and improved the recall rate from 62.5% to 73.3%, while the F-score also increased slightly by 6.9 % from 76.7% to 83.6%.

TABLE V: RESULTS OF CONVENTIONAL DNN MODELS (VARIOUS NEURONS)

| Model   | Accuracy | Precision | Recall | F-Score |
|---------|----------|-----------|--------|---------|
| DNN-64N | 0.931    | 0.972     | 0.733  | 0.836   |
| DNN-32N | 0.930    | 0.994     | 0.712  | 0.830   |
| DNN-16N | 0.909    | 0.993     | 0.625  | 0.767   |

Based on the accuracy metric, the DNN-3L model was the best model that reached the highest accuracy at 0.932 from the above experiments. As shown in Table IV and V, all DNN models performed well and improved the average accuracy to the highest of 0.926, whereas, the average accuracy of all machine-learning models only achieved 0.851, which was lower than the performance of deep learning models by 7.5%. Although the modified fuzzy decision tree improved the average accuracy to 0.888 for machine learning models, and the grey incidence analysis with reduced variables increased the average accuracy to 0.862, the overall performance of machine learning approaches still underperformed the deep learning approaches.

### C. Performance of the Feature Selections

To examine whether feature selection could increase the predictive accuracy for deep learning models, three feature selection methods were integrated with DNN models in the final experiment, including grey incidence analysis, Pearson correlation analysis and information gain. We only used the top 10 ranked variables as inputs from the results of feature selection to enable provably consistent configuration with the DNN models. By referring to Table VI, the average accuracy of DNN models with feature selection remained as higher over 0.930, and the accuracy of model DNN model with grey incidence was slightly higher than the other two models.

TABLE VI: RESULTS OF DNN MODELS WITH FEATURE SELECTION

| DNN (FS)            | Accuracy | Precision | Recall | F-Score |
|---------------------|----------|-----------|--------|---------|
| Grey Incidence      | 0.935    | 0.989     | 0.738  | 0.845   |
| Pearson Correlation | 0.934    | 0.994     | 0.729  | 0.841   |
| Information Gain    | 0.932    | 0.983     | 0.729  | 0.837   |

To conclude, the model accuracy could be improved in the range of 0.909 to 0.935 by using the deep neural network with various configurations. This result provided a good direction to aim for when developing the default prediction models. In addition, the feature selection methods reducing the dimensionality of input variables to include only 10 variables acquired the highest average accuracy than any other experiment results. However, all three strategies of feature selection achieved similar predictive accuracy, with slightly better performance being reported by the grey incidence analysis.

## IV. CONCLUSION

In this study, we first implemented four machine-learning approaches to predict the default cases, however, most models encountered challenges to resolve imbalance problem of the delinquency cases in our datasets. To improve the

predictive performance, we proposed another two strategies including grey incidence analysis and fuzzy decision tree to improve the predictive accuracy. The average accuracy were increased from 0.82 to 0.86 and 0.89 for each strategy respectively. The performance of DT model in machine learning approaches was improved by integrating the fuzzy partition with K-Means clustering method, and the predictive accuracy of the proposed FDT model was risen to higher value at 0.893.

Moreover, the input variables were rearranged for the grey incidence analysis to predict the default cases based on searching the geometrical similarity of data series. The predictive accuracy was increased from 0.806 to 0.891 if compared to its counterpart of KNN model using similar searching algorithm. Although the proposed two strategies integrated with machine learning approaches moderately increased the average accuracy to 4.5% and 7.2% respectively, we also implemented the newly developed deep learning approaches to examine whether the new techniques could improve the predictive accuracy.

The DNN models were evaluated with a different number of hidden layers and neurons, which also combined with the dropout layers and various activation functions to improve the training process. The experiment result indicated the DNN model with different number of layers increased the average accuracy to 0.930. On the other hand, the DNN models evaluated with different number of neurons suggested that the DNN-64N model achieved the highest accuracy of 0.931 if compared with other two models within the same group. Finally, the DNN models were evaluated with three different feature selection, and the experiment result reached the highest average accuracy of 0.934. The result also suggested that grey incidence analysis employed to rank variables achieved the highest accuracy of 0.935 and was considered as the best model if compared to the results of all experiments.

In conclusion, the experiment results showed the deep neural network performed better in most evaluation metrics and achieved an impressively high average accuracy of 0.930 if compare to the machine learning models which was 0.851. This research work could be extended to implement other deep learning algorithms to learn and trace the behaviors of the credit card holders and reduce the default risks for banking industries in the future.

## REFERENCES

- [1] J. Haslem and W. Longbrake, "A credit scoring model for commercial loans: Comment," *Journal of Money, Credit, and Banking*, vol. 4, no. 3, pp. 733-734, July 2015.
- [2] A. Steenackers and M. J. Goovaerts, "A credit scoring model for personal loans," *Insurance Mathematics Economic*, pp. 31-34, 1989.
- [3] T. Chou, "Apply two-stage grey incidence analysis with dynamic weights to improve the quality of default prediction," *Journal of Quality*, vol. 22, pp. 405-426, 2015.
- [4] I. Witten, E. Frank, M. Hall, and C. Pal, *Data Mining: Practical Machine Learning Tools and Techniques*, 3rd ed. Morgan Kaufmann, 2016.
- [5] A. Galathiya, A. Ganatra, and C. Bhensdadia, "Improved decision tree induction algorithm with feature selection, cross validation, model complexity and reduced error pruning," *International Journal of Computer Science and Information Technologies*, vol. 3, no. 2, pp. 3427-3431, 2012.
- [6] J. Alcalá-Fdez and J. Alonso, "A survey of fuzzy systems software: Taxonomy, current research trends and prospects," *IEEE Transactions on Fuzzy Systems*, vol. 24, no. 1, pp. 40-56, 2016.

- [7] L. Jouffe, "Fuzzy inference system learning by reinforcement method," *IEEE Transaction on System, Man, and Cybernetics*, vol. 28, pp. 338-355, 1998.
- [8] V. Levashenko, E. Zaitseva, and S. Puuronen, "Fuzzy classifier based on fuzzy decision tree," in *Proc. EUROCON 2007 International Conference on Computer As A Tool*, 2007, pp. 823-827.
- [9] V. Levashenko and E. Zaitseva, "Fuzzy decision trees in medical decision making support system, computer science and information systems," in *Proc. the 2012 Federated Conference on Computer Science and Information Systems*, 2012, pp. 213-219.
- [10] S. Guillaume and B. Charnomordic, "Learning interpretable fuzzy inference systems with fispro," *International Journal of Information Sciences*, vol. 181, no. 20, pp. 4409-4427, 2011.
- [11] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, and E. A. Fuad, "A survey of deep neural network architectures and their applications," *Neurocomputing*, vol. 234, pp. 11-26, 2017.
- [12] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Networks*, vol. 61, pp. 85-117, 2015.
- [13] D. Yamaguchi, G. Li, and M. Nagai, "Evaluation on the effectiveness of grey relational analysis models," in *Proc. the 11th Grey System Theory and Application Conference*, 2006, pp. 67-71.



mining.

**Tsungnan Chou** was born in Taiwan in 1963. He received the PhD degree in manufacturing engineering and operations management from the University of Nottingham in England in 1997. He is currently an assistant professor of finance at Chaoyang University of Technology where he joined in 1999. He has published more than 50 papers, both in conferences and journals. His research interests include artificial intelligence, machine learning and financial data



**Mingmin Lo** was born in Taiwan in 1974. She received the PhD degree with finance field in economics from the Central University of Finance and Economics in China in 2015. She is currently an assistant professor of finance at Chaoyang University of Technology where she joined in 2009. She has published more than 20 papers, both in conferences and journals. Her research interests include finance, investment and financial default issue.