Determination of Insurance Policy Using Neural Networks and Simplified Models with Factor Analysis Technique

YU-JU LIN[1], CHIN-SHENG HUANG[2], CHE-CHERN LIN[3]

1: Ph.D. Student, Department and Graduate Institute of Finance
National Yunlin University of Science and Technology &
Department of Finance, Fortune Institute of Technology

2: Department and Graduate Institute of Finance,
National Yunlin University of Science and Technology

3: Department of Software Engineering
National Kaohsiung Normal University

TAIWAN

kitty@center.fotech.edu.tw[1], huangcs@yuntech.edu.tw[2], cclin@nknucc.nknu.edu.tw[3]

Abstract: In this paper, we use feed forward neural networks with the back-propagation algorithm to build decision models for five insurances including life, annuity, health, accident, and investment-oriented insurances. Six features (variables) were selected for the inputs of the neural networks including age, sex, annual income, educational level, occupation, and risk preference. Three hundred insurants from an insurance company in Taiwan were used as examples for establishing the decision models. Six experiments were conducted in this study. These experiments were mainly categorized into two phases: Phase 1 (Experiments 1 to 3) and Phase 2 (Experiments 4 to 6). In Phase 1, we used the six features as the inputs of the neural networks. In Phase 2, we employed the factor analysis method to select three more important features from the six features. In Phase 1, Experiment 1 used a single neural network to classify the five insurances simultaneously while Experimental 2 utilized five neural networks to classify them independently. Experiments 1 and 2 adopted the purchase records of primary and additional insurances as experimental data. Experiment 3, however, utilized the primary insurance purchase dada only. In Phase 2, we repeated the similar experimental procedure as Phase 1. We also applied statistical methods to test the differences of the classification results between Phases 1 and 2. Discussion and concluding remarks are finally provided at the end of this paper.

Key-words: - Insurance policy, Neural networks, Back-propagation algorithm, Classification, Factor analysis, Feature extraction.

1. Introduction

The determination of the best insurance policy for potential insurance buyers is a big challenge for insurance consultants. Customers come to buy insurances with different backgrounds, economic conditions, risk preference degrees, and needs. In general, there are five major types of insurances in the insurance industry including life, annuity, health, accident, and investment-oriented insurances. The purpose of buying insurance, the amount and the period of benefit payments returned, and the coverage of insurance are quite different among the five insurances. Building a decision support system for the insurance consultants to determine an appropriate insurance for their customers now becomes one of important issues in the insurance industry. Some studies used statistical methods to build the decision models for choosing insurance policy such as regression models, k-mean clustering techniques, discriminate analyses, etc. In addition, some researches employed artificial intelligence techniques to establish classifiers for determining insurance policy. In artificial intelligence, two methodologies are often used to establish decision models: the methodology of using human experts (expert systems) and the methodology of using data (data miming). Experts systems are used to build decision support systems if human experts are available to help people build the knowledge base (rule base) for the systems. Fuzzy expert systems are one of popular expert system techniques in solving real-world problems with successful applications. Basically, the fuzzy expert system technique mimics the verbal expressions and thinking process of human beings. It is suitable for solving uncertainty problems using fuzzy inference process. Membership functions are used to represent these uncertainty situations. The think processes are then expressed in fuzzy rules. In general, there are three major components in a fuzzy expert system including inference engine conducting fuzzy output, user interfaces allowing users to compose the fuzzy variables and fuzzy rules, and rule base storing fuzzy rules. Contrary to expert systems, data mining employs existing data to build the decision models without humane experts. The data mining approaches include neural networks, decision trees, basket analyses, etc. In addition to the techniques of expert systems and data mining, genetic algorithms are utilized to build decision models as well. A genetic algorithm is basically an optimization algorithm to find the optimal solution in a very large searching space. Based on the Darwin's evolution theory, a genetic algorithm employs evolution processes to reach the optimal solution including crossover, selection, and mutation. Fitness functions are utilized to measure the performance of evolution results. Genes and chromosomes are also used to determine fitness functions in genetic algorithms. Neural networks are one of popular techniques to for classification problems. In this paper, we use networks to generate the decision models to determine insurance policy. In Section 3, we describe the computational details about the neural networks used in this paper. Below, we briefly introduce the characteristics of the five insurances.

Like a contract between an insurance buyer (insurant) and seller (insurer), the life insurance specifies insured amount for the buyer and seller when the insurant purchases the life insurance. The insurant receives a pre-specified amount whenever the pre-set conditions for benefit payments are met. The situation (dead or alive) of an insurant is the condition of determining the benefit payment. For annuity insurance, the policy holders of this insurance will be periodically provided benefit payments in their life time. The annuity insurance is therefore to be one of popular alternatives for retirement programs for those people who want to combine insurance together with retirement stipends. The purpose of health insurance is to make up for medical expenditures and to get the benefit to compensate for the reduction of income when the insurant is sick. The purpose of accident insurance is to cover the loss of accident. Therefore, accident insurants will receive a pre-specified amount when they get hurt due to an accident. As a new kind of insurance in the insurance market, the investment-oriented insurance provides not only the functionality of insurance but also the purpose of investment. In investment-oriented insurance, the insurants are responsible for the investment risks, and the insured amounts are determined by the earnings of the invested targets.

In this paper, we used feed forward neural networks with the back-propagation algorithm to build decision models for the five insurances mentioned above. Six features were selected for the inputs of the neural networks including age, sex, annual income, educational level, occupation, and risk preference. Three hundred insurants from an insurance company in Taiwan were used as examples for the neural networks. Six experiments were conducted in this study. These experiments were mainly categorized into two phases: Phase 1 (Experiments 1 to 3) and Phase 2 (Experiments 4 to 6). In Phase 1, we used the six features as the inputs of the neural networks. In Phase 2, we employed the factor analysis method to select three more important features from the six features. In Phase 1, Experiment 1 used a single neural network to classify the five insurances simultaneously while Experimental 2 utilized five neural networks to classify them independently. Experiments 1 and 2 adopted the purchase records of primary and additional insurances as experimental data. Experiment 3, however, utilized the primary insurance purchase dada only. In Phase 2, we repeated the similar experimental procedure as Phase 1. We also applied statistical methods to test the differences of the classification results between Phases 1 and 2. Discussion and concluding remarks are finally provided at the end of this paper.

2. Related works

Conceptually, neural networks emulate the brain structure and the recognition process of human beings. Many neural network models have been developed to solve real world problems including feed forward neural network with back-propagation algorithm, self-organizing map, winner-take-all network, etc [1]. Basically, a neural network is hierarchically structured with layers. Each of layer consists of several nodes. Weighs are used to connect two different nodes in adjacent layers. Outputs are obtained by a sequence of mappings from inputs. Training algorithms are employed to update the weights to get desired outputs. Normally, training algorithms spend a lot of computational time in updating weights. Some researches discussed the implementation feasibility of neural networks where some different kinds of decision regions have been proved to be implemented by multi-layered neural networks without any training algorithms [2-5]. Mastorakis proposed an optimal multi-layer structure for back-propagation networks where based on Bayesian Information Criterion (BIC), the complexity of a neural network can be gradually increased until an acceptable error [6].

In practice, to perform neural network computing from a data set, the inputs of the neural network are the features extracted form the original data set. Extraction methods are commonly used to extract the features from the original data set by some necessary computations before using neural networks. Some of feature extraction methods decompose original data into orthogonal components such as Radial-Basis Function [7] and wavelet function [8]. Neural networks can be combined by other soft computing methods such as fuzzy logic [9] and genetic algorithm [10].

Insurance business is one of important industries in a country. Some research methodologies applied statistical methods to insurance-related studies. Recently with the rapid promotion of computer technologies, a lot of researches then used the artificial intelligence techniques (or so-called soft computing techniques) to solve the problems regarding insurance. Shapiro indicated that three artificial intelligence techniques have often been used in insurance-related researches including neural networks, fuzzy logic, and genetic algorithms [11]. He further pointed out how the three soft computing techniques were utilized to solve the insurance-related problems. [11]. Below, based on [11], we summarize the previous works related to the neural-network-based researches on insurance.

The feed-forward neural network model with the back-propagation training algorithm has been utilized to establish underwriting bond [11,12]. Brockett et al. employed Kohonen self-organized feature map to detect claim frauds for body injury in car accidents [11,13]. Viaene et al. used Bayesian learning neural networks to detect the claim fraud for automobile insurance [11,14]. In addition, several studies applied neural networks to predict early financial crisis or bankrupts for insurance companies [11,15,16]. Meanwhile, some of previous studies focused on the topic of medical-care-related insurances [11]. For example, Ismael presented the prediction model for the mortality of acute myocardial infarction in-hospital patients using neural networks [11,17]. Saemundsson discussed the application of neural networks on dental care insurance [11, 18].

In addition to the above bankrupt predictions on the insurance companies, several studies used neural networks to predict bankrupts or financial crises for the companies of other industries [19-26].

3. The back-propagation algorithm

The network architecture for the back-propagation algorithm is a feed forward neural network with a couple of layers. Each of layers contains several nodes. Weights are employed to connect nodes in adjacent layers. There are two phases during employing the back-propagation algorithm: a forward computing phase and a backward updating phase. During the forward computing phase, inputs are mapped using activation functions (discussed later) to produce their corresponding outputs.

Figure 1 shows an example of a two-layer feed forward neural network. It contains two inputs which form a two-dimensional input space, three nodes in the first layer, and one output node in the second layer. In the figure, each node in the network structure uses an activation function to map the input to its corresponding output. We explain the mapping procedure of a node in a neural network by a simple example. Consider the output node y in the second layer (current layer) in Figure 1. We first generate the input of node y. The input of a node y in the second layer is the summation of the products of the outputs of all nodes in the previous layer (z_1 to z_3 in the first layer), and their corresponding weight (w_1 to w_3). A general equation to describe the above summation is given by

$$\text{input of } y = \sum_{i=1}^{n} w_i z_i \tag{1}$$

where n (in Figure 1, n=3) is the number of nodes in previous layer, z_i is a node in the previous layer, and w_i is the weigh connection z_i to node y.

After obtaining the input of y using Eq. 1, an activation function is then used to perform the mapping from the input to the corresponding output for node y. The sigmoid function is one of popular activation functions used in neural networks. It is given by

$$sigmoid\ (t) = \frac{1}{1+e^{-t}} \tag{2}$$

Figure 2 shows the mapping diagram of the sigmoid function. The sigmoid function maps from the domain of (-∞, ∞) to the range of (0, 1) with a smooth transition.

During the updating phase, weights are updated from outputs to inputs backwardly using updating formulas. The back-propagation algorithm is a popular updating algorithm used in feed forward neural networks. It basically uses the gradient descent method to optimize the weights. The gradient descent method optimizes the weights according to the square errors between the computed outputs and the desired outputs.

The procedure of the back-propagation algorithm can be summarized as follows [1, 27]:

Step 1: Design the neural work structure to solve the problem: The number of the inputs for the neural network is equal to the number of attributes of the problem. The number of outputs is also equal to the number of outputs for problem. The number of layers and the number of nodes in each layer are normally determined by research designers. In general, the number of layers for a back-propagation neural network is two (one hidden layer).

Step 2: Initialize all weight in the neural network with small random numbers.

Step 3: Forwardly calculate the outputs of the neural network from its inputs using Eqs. (1) and (2).

Step 4: Update the weights by

$$w_{ij}^{k+1} = w_{ij}^{k} + \delta\phi_j x_i \tag{3}$$

where i represents the i^{th} node in the pervious layer (in the forward direction; refer to Figure 1) and j represents the j^{th} node in the current layer; w_{ij} is the weight linking node i to node j; k is an iterative indicator; δ is a step size ($0 < \delta < 1$); x_i is the output of node i; ϕ_j is given by

$$\begin{aligned} \phi_j &= y_j(1-y_i)(y'_j - y_j) \text{ if } j \text{ is an output node} \\ &= x_j(1-x_j)\sum_l \phi_l w_{jl}) \text{ if } j \text{ is a hidden node} \end{aligned} \tag{4}$$

where y_j is the computed output of node j, y'_j is the desired output of node j, and l is the number of nodes in the previous layer (in the backward direction; refer to Figure 1) .

Step 5: Terminate the iterations if the pre-defined iteration number is reached or the error criterion is met; otherwise, go to Step 3.

4. Experiments and Results

4.1 Phase 1: Using six variables (features)

The data were collected from an insurance company in Taiwan. Three hundred insurants were selected as the samples for the experiments. These insurants might purchase multiple insurances (primary and additional insurances) simultaneously. Six input variables were used in the experiments including age, sex, annual income, educational level, occupation, and risk preference. All of them were encoded in integer. The details of the input variables are described as follows:

- Age: encoded in year.
- Sex: categorical, 1 for male; 2 for female.
- Annual income: Encoded by a unit of 10,000 NTDs (New Taiwan Dollars).
- Educational level: categorical with 9 values:
  1: Elementary school;
  2: Junior high school;
  3: Senior high school;
  4: Vocational high school;
  5: Junior college or community college;
  6: Technical college;
  7: University;
  8: Master degree;
  9: Doctoral degree.
- Occupation: categorical, ranked in the ascending order of the occupational risks from 1 (lowest) to 6 (highest).
- Risk preference: categorical with values from 1 (lowest risk preference) to 10 (highest risk preference).

Table 1 shows the frequencies of insurance purchases (including primary and additional insurances) for the 300 insurants.

In Phase 1, three experiments were conducted to build the classification models for the five insurances. Experiment 1 used a single neural network with five outputs to classify the five insurances simultaneously. Experiment 2 used five neural networks (each with one output) to classify the five insurances independently. Experiments 1 and 2 used the purchase records of primary and additional insurances as experimental data. Experiment 3 used the data of primary insurances only. Below, we explain the three experiments.

Experiment 1:

In this experiment, we used a single two-layer neural network to integrate all of the five insurances simultaneously. The neural network contained 6 inputs, 10 nodes in the first layer, and 5 output nodes in the second layer. Each of the output nodes was associated with a single insurance.

We used three rounds in this experiment. In the first round, all of the samples were used as training examples for building the decision models. In second round, we used 2/3 of samples as a training set and 1/3 as a test set. In the last round, we used 3/4 of samples as a training set and 1/4 as a test set. The classification accuracies of Experiment 1 are shown in Table 2

Experiment 2:

In this experiment, we used five two-layer neural networks to classify the five insurances independently. Each of the five neural networks contained 6 inputs, 10 nodes in the first layer, and one output node in the second layer.

Again we used the same procedure as Experiment 1 to conduct three rounds in this experiment. The classification accuracies of Experiment 2 are shown in Table 2

Experiment 3:
In this experiment, we only considered those insurants who purchased primary insurances as the samples. Of the five insurances, the data of health insurance and accident insurance are different from the data used in Experiments 1 and 2. The sample sizes of health insurance and accident insurance for Experiment 3 are 60 and 22, respectively. The samples of the rest of the insurances (life, annuity, and investment-oriented insurances) remain the same. Similarly, in this experiment, we repeated the same three rounds as Experiments 1 and 2. Table 3 shows the classification accuracies of Experiment 3.

4.2 Phase 2: Simplified models using three more important variables (features)

The complication of neural networks depends on their structures. More nodes cause more computations. In this study, we first used six features as the input variables. To simplify the neural network structures, we used the factor analysis technique to reduce the input dimensions (the number of input variables). The factor analysis technique is a statistical method to reduce data sizes by removing less important variables from original data based on their importance. Eigenvalues are used to measure the importance of the input variables (higher eigenvalues mean more important). This is called feature extraction procedure. Popular extraction techniques include principal components, unweighted least squares, generalized least squares, maximum likelihood, principal axis factoring, and alpha factoring methods [28]. Two common criteria are utilized to extract more important features from the original data. The first one is to set the threshold of eigenvalue to be 1. The second one is to select features using a scree plot where the eigenvalues of the input data are plotted in a descending order. One normally sets the threshold of the eigenvalues to be the point where a sharp change (decline) occurs. In addition, a loading plot depicts the visual expression of the loading of the original variables on the dimensions (components) selected by the factor analysis method. Meanwhile, in some cases, one uses rotations to perform linear mapping to get better interpretation of data. Commonly used rotation methods include varimax, direct oblimin, quartimax, equamax, and promax methods [28].

The factor analysis technique is widely applied in determining the dimensions in a questionnaire. In this study, we apply it to reduce the number of input variables. The factor analysis technique has been utilized in many areas with successful applications [29-31].

The original input variables include age, sex, annual income, educational level, occupation, and risk preference. Table 4 displays the descriptive statistics of the six variables. Figure 3 shows the scree plot of the eigenvalues of the input variables. If we set the threshold to be 1, three variables are selected. Table 5 displays the results of factor analysis. Based on Table 5, three components (input variables) were selected as the inputs of the simplified models including education, occupation, and age. Table 6 demonstrates the total variance explained by the three selected components. From the table, one finds that 71.304 % of accumulated variance can be explained by the three components. Figure 4 demonstrates the loading plot.

In this section, we used the principal component analysis method as the extraction method and the varimax method as the rotation method. We repeated the experimental procedure in Section 4.1. Tables 7 and 8 demonstrate the experimental results.

4.3 Statistical tests

As mentioned early, the computational complexities of neural networks are dependent on the network structures. More nodes and layers in a neural network might cause more computations. In this study we applied statistical methods to analyze the classification performance between the neural networks having six input variables (refer to Phase 1 in Section 4.1) and the neural networks having three input variables (refer to Phase 2 in Section 4.2). The statistical hypothesis is shown as follows:

H: The classification accuracies are different between the neural networks having the six input variables and three input variables.

The statistical procedure for the hypothesis is described as follows

Step 1: Test the correlation between the neural networks having the six input variables and three input variables using Pearson product-moment correlation method.

Step 2: Use the paired samples t-test.

Table 9 shows the Pearson correlation coefficient obtained by Step1. The correlation is significant at the 0.001 level with a two-tailed significance test. This result supports us to perform the test in Step 2. Table 10 shows the results of paired samples statistics. Table 11 displays the results of paired samples test on the pair of 6 variables – 3 variables. From Table 11, we accept the hypothesis and conclude that the classification accuracies are different between the neural networks having the six input variables and three input variables.

5. Discussion

According to the experimental results shown in Tables 2 and 3, in general, the classification accuracies in Experiment 2 are better than those in Experiment 1. Furthermore, the classification accuracies in Experiment 3 are better than those in Experiment 2. Below, we discuss the experimental results in details.

In Experiment 1, we considered all of the five insurances as a whole and integrated the five insurances into a single neural network. In Round 1 of this experiment, the accuracy for the overall samples is 64.6%. It is lower than the results in Experiment 2. In Round 2 and Round 3, there existed over-fitting problems since the accuracies of the training set are better than those of the test sets. The classification performance of Round 2 is better than that of Round 3.

In Experiment 2, if we used all of examples as a training set (Round 1), the classification accuracies of the five insurances are all higher than 74%. Of the five accuracies, the accuracies of annuity insurance and health insurance are higher than 89%. In Round 2 and Round 3 of the experiment, we divided the samples into two sets (training set and test set) according to different size ratios of training sets and test sets. Over-fitting problems still occur in life, investment-oriented, and accident (Round 3 only) insurances.

In Experiment 3, we only took the insurants who purchased the primary insurances as experimental examples. The sample sizes of health insurance and accident insurance decreased. The experimental results show that the classification accuracies of the two insurances are better than those in Experiment 2

In Experiments 1 and 2, we used the insurants who purchased the primary and additional insurances as experimental examples. However, in insurance practice, the actual value of the additional insurance is much less important than that of the primary one. That might be the reason why the accuracies in Experiment 3 are better than those in Experiment 2.

The experimental results show that the classification results of using five independent neural networks to classify the five insurances independently are better than those of using one single neural network to classify the five insurances simultaneously. Meanwhile, the classification accuracies of using the examples of primary insurances purchases are better than those of using the examples of primary and additional insurance purchases.

The experiments in Phase 1 used the original six variables as the inputs of the neural networks. The experiments in Phase 2 used the three variables whose eigenvalues are higher than 1 as the inputs of the neural networks. We employed the statistical methods to test if the experimental results between Phase 1 and Phases 2 are different. Based the results of the statistical test, we conclude that the classification accuracies of the neural networks using the six input variables are better than those of using the three input variables whose eigenvalues are higher than 1.

6. Conclusions

We used feed forward neural networks to build decision models for five insurances including life, annuity, health, accident, and investment-oriented insurances. Six features were selected for the inputs of the neural networks including age, sex, annual income, educational level, occupation, and risk preference. Three hundred insurants from an insurance company in Taiwan were used as examples for establishing the decision models.

Six experiments were conducted in this study. These experiments were mainly categorized into two phases: Phase 1 (including Experiments 1 to 3) and Phase 2 (including Experiments 4 to 6). In Phase 1, we used the original six features as the inputs of the neural networks. In Phase 2, from the original six features, we used 3 more important features determined by the factor analysis method where the threshold of eigenvalue for selecting the features is set to 1.

In Phase 1, Experiment 1 used a single neural network with five output nodes to classify the five insurances simultaneously. Experimental 2 improves the classification accuracies of Experiment 1 by using five neural networks (each with one output) to classify the five insurances independently. Experiments 1 and 2 used the purchase records of primary and additional insurances as experimental data. Experiment 3 further improves the classification accuracies of Experiment 2 by using the data of primary insurances purchases only.

In Phase 2, we used three features as the inputs of the neural networks including age, occupation, and education. We repeated the same experimental procedure as Phase 1 and conducted three experiments (Experiments 4 to 6). The experimental setup and process of Experiment 4 is similar to those of Experiment 1 in Phase 1, except for the inputs of the neural network. Consequently, Experiment 5 is similar to Experiment 2, and Experiment 6 is similar to Experiment 3. In addition, we employed the statistical methods to test if the experimental results of Phase 1 and Phase 2 were different.

Based on the experimental results, the suggestions for building the classifiers to determine insurance policy are drawn as follows:

- Using five independent neural networks to classify the five insurances independently is better than using one single neural network to classify the five insurances simultaneously.
- Using the data of primary insurance purchases is better than using the data of primary and additional insurance purchases.
- Using the neural networks with the six original input variables is better than using those with three input variables (education, occupation, and age).

As for the directions of the future studies, it might be a good idea to use more samples from different insurance companies to build more general decision models for determining the purchase of insurances.

References:

[1] R.P. Lippmann, An introduction to computing with neural nets, *IEEE Acoustic, Speech, and Signal Processing ASSP Magazine*, April, 1987, pp. 4-12.

[2] C. Lin, A constructive algorithm to implement celled decision regions using two-layer perceptrons, *WSEAS Transaction on Information Science and Applications*, Issue 9, Vol. 3, 2006, pp. 1154-1160.

[3] C. Lin, Partitioning capabilities of multi-layer perceptrons on nested rectangular decision regions part I: algorithm, *WSEAS Transaction on Information Science and Applications*, Issue 9, Vol. 3, 2006, pp. 1674-1680.

[4] C. Lin, Partitioning capabilities of multi-layer perceptrons on nested rectangular decision regions part II: Properties and Feasibility, *WSEAS Transaction on Information Science and Applications*, Issue 9, Vol. 3, 2006, pp. 1681-1687.

[5] C. Lin, Partitioning capabilities of multi-layer perceptrons on nested rectangular decision regions part III: Applications, *WSEAS Transaction on Information Science and Applications*, Issue 9, Vol. 3, 2006, pp. 1688-1694.

[6] N.E. Mastorakis, The optimal multi-layer structure of backpropagation networks, *WSEAS Transaction on Information Science and Applications*, Issue 9, Vol. 3, 2006, pp. 1632-1637.

[7] Q. Zhu, Y. Cai, L. Liu, A global learning algorithm for a RBF network, *Neural Networks*, Vol. 12, 1999, pp. 527-540.

[8] A. Subasi, M. Yilmaz, H.R. Ozcalik, Classification of EMG signals using wavelet neural netwrok, *Journal of Neuroscience Methods*, 2006, Vol. 156, pp. 360-367.

[9] M. Pasquier, C. Quek, M Toh, Fuzzylot: a novel self-organising fuzzy-neural rule-based pilot system for automated vehicles, *Neural Networks*, Vol. 14, 2001, pp. 1099-1112.

[10] Y.A. Alsultanny, M.M. Aqel,, Pattern recognition using multilayer neural-genetic algorithm, *Neurocomputing*, Vol. 51, 2003, pp. 237-247.

[11] A.F. Shapiro, The merging of neural networks, fuzzy logic, and genetic algorithms, *Insurance: Mathematics and Economics*, Vol. 31, 2002, pp. 115-131.

[12] M.L. Vaughn, E. Ong, S.J. Cavill, Interpretation and knowledge discovery from a multilayer perceptron network that performs whole life assurance risk assessment, *Neural Computing and Applications*, Vol. 6, 1997, pp. 201–213.

[13] P.L. Brockett, X. Xia, R.A. Derrig, Using Kohonen's self-organizing feature map to uncover automobile bodily injury claims fraud, *JRI*, Vol. 65 No. 2, 1998, pp. 245–274.

[14] S. Viaene, G. Dedene, R.A. Derrig, Auto claim fraud detection using Bayesian learning neural networks, *Expert Systems with Applications*, Vol. 29, 2005, pp. 653-666.

[15] P.L. Brockett, W.W. Cooper, L.L. Golden, U. Pitaktong, A neural network method for obtaining an early warning of insurer insolvency, *Journal of Risk and Insurance*, Vol. 61, No. 3, 1994, pp. 402.

[16] C. Huang, R.E. Dorsey, M.A. Boose, Life insurer financial distress prediction: a neural network model, *Journal of Insurance* Regulation Winter Vol. 13, No. 2, 1994, pp. 131–167.

[17] M.B. Ismael, Prediction of mortality and

in-hospital complications for acute myocardial infarction patients using artificial neural networks, Ph.D. Dissertation, *Duke University*, 1999.

[18] S.R. Saemundsson, Dental caries prediction by clinicians and neural networks, Ph.D. Dissertation, *University of North Carolina*, Chapel Hill, 1996.

[19] K. Lee, D. Booth, P. Alam, A comparison of supervised and unsupervised neural networks in predicting bankruptcy of Korean firms, *Expert Systems with Applications*, Vol. 29, 2005, pp. 1–6.

[20] H. Jo, I. Han, H. Lee, Bankruptcy prediction using case-based reasoning, neural networks, and discriminant analysis, *Expert Systems with Applications*, Vol. 13, 1997, pp. 97-108.

[21] K. Kiviluoto, Predicting bankruptcies with self organizing map, *Neurocomputing*, Vol. 21, 1998, pp. 191–201.

[22] K.C. Lee, I. Han, Y. Kwon, Hybrid neural network models for bankruptcy predictions, *Decision Support Systems*, Vol. 18, 1996, pp. 63–72.

[23] M. Leshno, Y. Spector, Neural network prediction analysis: The bankruptcy case, *Neurocomputing*, Vol. 10, 1996, pp. 125–147.

[24] C. Serrano-Cinca, Self organizing neural networks for financial diagnosis, *Decision Support Systems*, Vol. 17, 1996, pp. 227–238.

[25] Z.R. Yang, M.B. Platt, H.D. Platt, Probability neural network in bankruptcy prediction, Journal of Business Research, Vol. 44, 1999, pp. 67–74.

[26] G. Zhang, M.Y. Hu, B.E. Patuwo, D.C. Indro, Artificial neural networks and bankruptcy prediction general framework and cross-validated analysis, *European Journal of Operational Research*, Vol. 116, 1999, pp. 16–32.

[27] C. Lin, Algorithms to determine the feasibilities and weights of multi-layer perceptrons with application to speech classification, Ph.D. Dissertation, University of Pittsburgh, PA, USA, 1996.

[28] SPSS Inc., *On-line documentation: SPSS 13.0 for Windows*, 2004.

[29] A. Marusic, Factor analysis of risk for coronary heart disease: an independent replication, *International Journal of Cardiology*, Vol. 75, 2000, pp. 233-238.

[30] N. Vukasinovic, J. Moll, N. Kunzi, Factor analysis for evaluating relationships between herd life and type traits in Swiss Brown cattle, *Livestock Production Science*, Vol. 49, 1997, pp. 227-234.

[31] H. Ninomiya, T. Onitsuka, C. Chen, N. Kinukawa, Possible overlapping potentials of the auditory P50 in humans: factor analysis of middle latency auditory evoked potentials, *Electroencephalography and Clinical Neurophysiology*, Vol. 104, 1997, pp 23-30.

Table 1: Insurance purchases for the five insurance

life	annuity	Health	accident	Investment -oriented
143	31	182	169	156

Table 2: Classification accuracies of Experiments 1 and 2 (in percentage)

Round		Integration of the five insurances (Experiment 1)	Individual Insurances (Experiment 2)				
			Life	Annuity	Health	Accident	Investment -oriented
1	Overall samples	64.6	85.6	89.6	91.3	75	74.6
2	Training (2/3 of samples)	80	96.05	81.5	75	51.5	86
	Test (1/3 of samples)	30	49	77	58	60	53
3	Training (3/4 of samples)	53	95.1	88.8	49.7	94.2	87
	Test (1/4 of samples)	18.6	42.6	92	66.7	44	52

Table 3: Classification accuracies of Experiment 3 (in percentage)

Round		Integration of the five insurances	Individual Insurances	
			health	accident
1	Overall samples	61	85.6	97.6
2	Training (2/3)	64	96	95.8
	Test (1/3)	26	75	86
3	Training (3/4)	77.3	90.2	95.5
	Test (1/4)	33.3	74.6	84

Table 4: Descriptive statistics

Variable	Mean	Std. Deviation	N
Age	38.293	10.388	300
Sex	1.54	0.499	300
Income	69.157	43.940	300
Education	5.33	1.724	300
Risk Pref.	5.041	2.125	300
Occupation	1.37	0.703	300

Table 5: The result of factor analysis

Variable	Component		
	1	2	3
Age	-.374	.007	.843
Sex	-.301	-.757	-.166
Income	.508	.078	.731
Education	.769	-.189	-.164
Risk Pref.	.694	.143	.065
Occupation	-.271	.834	-.088

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

Table 6: Variance explained

Component	% of Variance	Cumulative %
1	27.400	27.400
2	24.529	51.929
3	19.375	71.304
Extraction Method: Principal Component Analysis.		

Table 7: Classification accuracies of Experiments 4 and 5 of Phase 2: simplified models (%)

Round		Integration of the five insurances (Experiment 1)	Individual Insurances (Experiment 2)				
			Life	Annuity	Health	Accident	Investment -oriented
1	Overall samples	28.5	65.7	91.3	71.3	73.3	65.0
2	Training (2/3 of samples)	44.5	77.5	94.0	75.5	77.5	74.5
	Test (1/3 of samples)	20.0	49.0	85.0	70.0	52.0	48.0
3	Training (3/4 of samples)	39.0	67.1	92.4	73.3	74.6	69.8
	Test (1/4 of samples)	18.6	46.7	86.7	65.3	57.3	48.0

Table 8: Classification accuracies of Experiment 6 of Phase 2: simplified models (%)

Round		Integration of the five insurances	Individual Insurances	
			health	accident
1	Overall samples	61.3	70.0	93.7
2	Training (2/3)	47.5	87.0	94.0
	Test (1/3)	20.0	67.0	90.0
3	Training (3/4)	57.0	87.5	91.5
	Test (1/4)	18.6	69.3	86.6

Table 9: The results of paired samples Correlations

Pair	N	Correlation	Sig.
6 Var. – 3 Var.	45	0.817	.000

Table 10: The results of paired samples statistics

Pair	Mean	N	Std. Deviation	Std. Error Mean
6 Var.	71.19	45	21.35	3.18
3 Var.	65.387	45	21.87	3.260

Table 11: The results of paired samples test

Paired Differences			t	Sig. (2-tailed)
Mean	Std. Devi.	Std. Error Mean		
5.800	13.089	1.951	2.972	0.005.

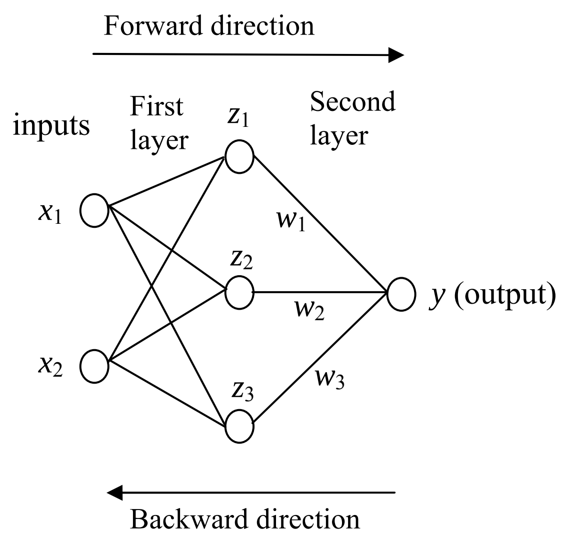

Figure 1: A two-layer feed forward neural network

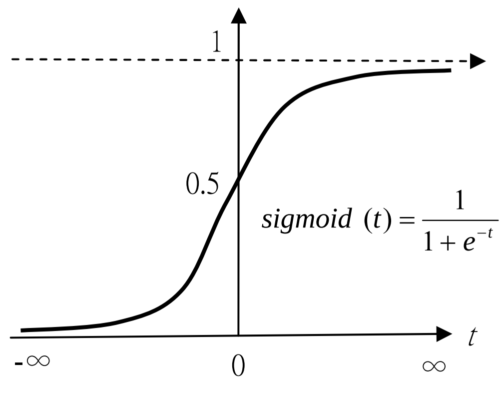

Figure 2: The sigmoid function

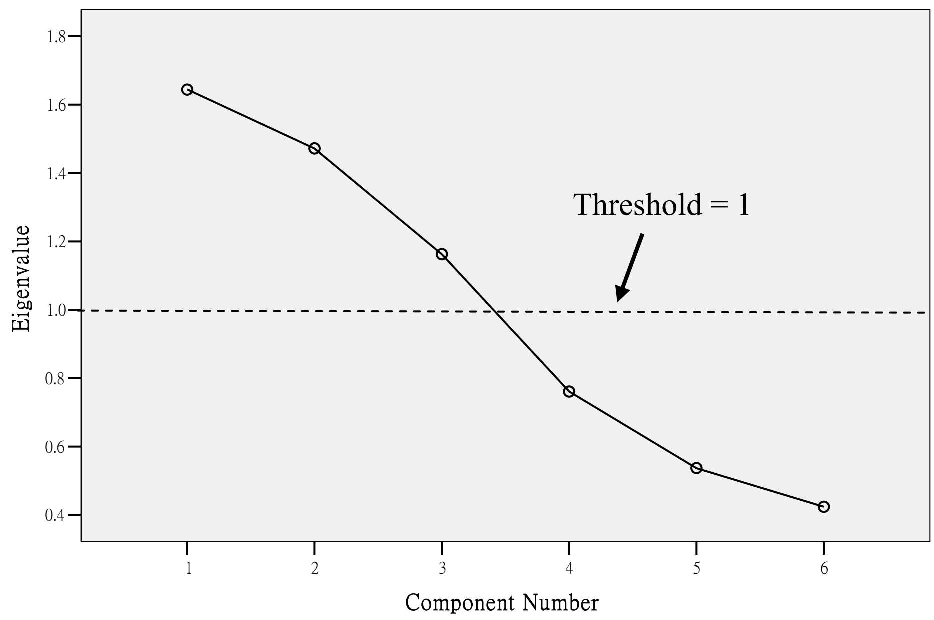

Figure 3: The scree plot

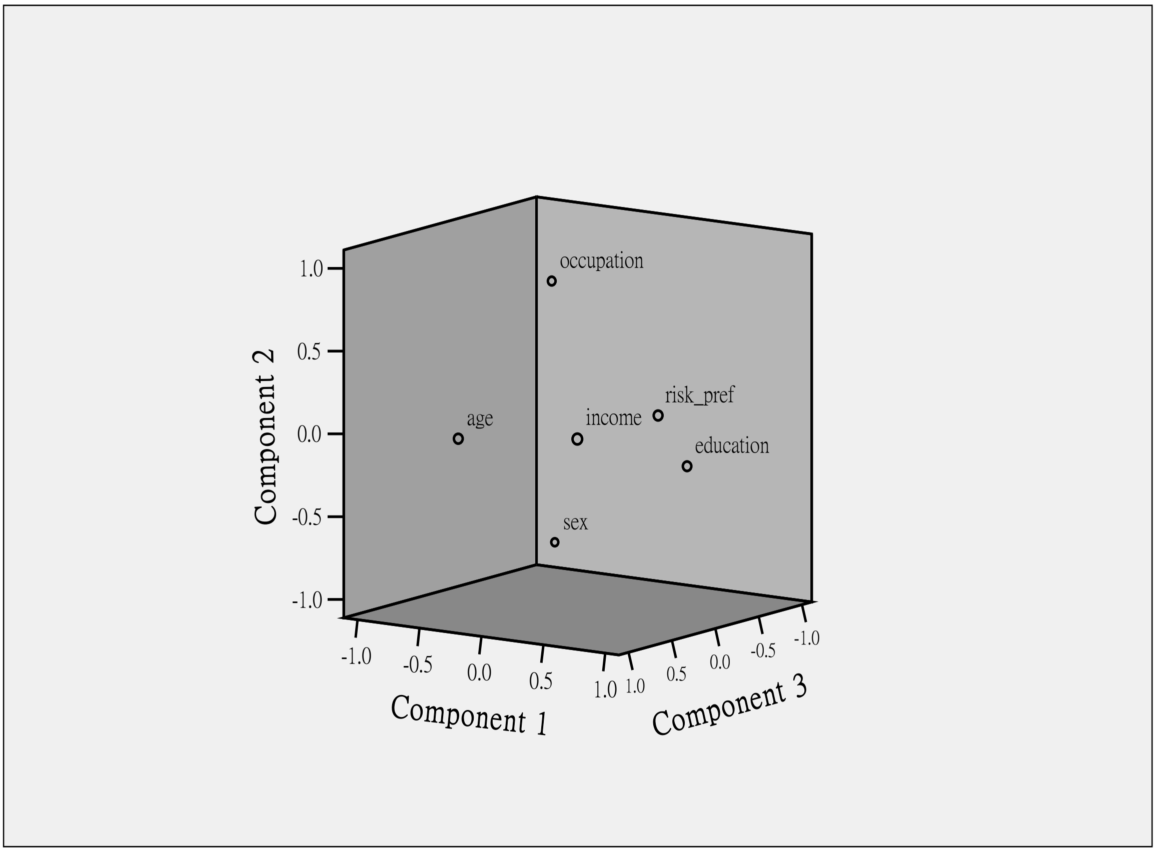

Figure 4: The loading plot