Evaluation Models for Choosing Insurance Policy Using Neural Networks

CHIN-SHENG HUANG¹, YU-JU LIN², CHE-CHERN LIN³

¹: Department and Graduate Institute of Finance, National Yunlin University of Science and Technology
²: Ph.D. Student, Department and Graduate Institute of Finance National Yunlin University of Science and Technology & Department of Finance, Fortune Institute of Technology
³: Department of Software Engineering National Kaohsiung Normal University TAIWAN

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 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. There experiments were conducted in this study. The first one considered the five insurances as a whole and established a single neural network integrating all of the five insurances in its structure. The second one built five individual neural networks independently for the five insurances, respectively. The first two experiments used the purchase records of primary and additional insurances as experimental data. The last experiment only used the data of primary insurances. We showed the experimental results and discussed problematic issues on the experiments. Based on the experimental results, the suggestions for build classifiers for insurance purchase are drawn as follows: (1) 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; (2) using the data of primary insurances purchases is better than using the data of primary and additional insurance purchases. Finally, the possible directions for future studies are provided at the end of this paper.

Key-words: - Insurance policy; Decision making; Neural Networks, Back-propagation algorithm; Classification

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 live insurance specifies insured amount for the buyer and seller when the insurant purchases the live insurance. The insurant receives a pre-specified amount whenever the pre-set conditions for benefit payments are met. The situation (dead or live) 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 use 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. Three experiments were conducted in this study. The first one considered the five insurances as a whole and used a single neural network to integrate all of the five insurances in its structure. The second one built five individual neural networks independently for the five insurance, respectively. The first two experiments used the purchase records of primary and additional insurances as the examples for building the decision models. The last one utilized the purchase data of primary insurances only. We analyzed the experimental results and discussed the problematic issues related to the experiments. Concluding remarks are finally provided at the end of this paper.

2. Related works

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 [1]. He further pointed out how the three soft computing techniques were utilized to solve the insurance-related problems. [1]. Below, based on [1], we summarize the previous works related to the neural-network-based researches on insurance.

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 [2]. The feed-forward neural network model with the back-propagation training algorithm has been utilized to establish underwriting bond [1,3]. Brockett et al. employed Kohonen self-organized feature map to detect claim frauds for body injury in car accidents [1,4]. Viaene et al. used Bayesian learning neural networks to detect the claim fraud for automobile insurance [1,5]. In addition, several studies applied neural networks to predict early financial crisis or bankrupts for insurance companies [1,6,7]. Meanwhile, some of previous studies focused on the topic of medical-care-related insurances [1]. For example, Ismael presented the prediction model for the mortality of acute myocardial infarction in-hospital patients using neural networks [1,8]. Saemundson discussed the application of neural networks on dental care insurance [1, 9].

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 [10-17].

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

$$\text{input of } y = \sum_{i=1}^{n} w_i z_i$$  \hspace{1cm} (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 weight 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

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

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 [2]:

**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

\[ w_{ij}^{k+1} = w_{ij}^{k} + \delta \phi_{j} x_i \]  

(3)

where \( i \) represents the \( i \)th node in the previous 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; \( \delta \) is a step size (0 < \( \delta \) < 1); \( x_i \) is the output of node \( i \); \( \phi_j \) is given by

\[ \phi_j = y_j (1-y_j) (y'_j - y_j) \]  

for output node

\[ \phi_j = x_j (1-x_j) \sum_i \phi_i w_{ij} \]  

for hidden node

(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. Experimental Design and Results

4.1 Data collection

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.

<table>
<thead>
<tr>
<th>Life annuity</th>
<th>Health</th>
<th>Accident</th>
<th>Investment-oriented</th>
</tr>
</thead>
<tbody>
<tr>
<td>143</td>
<td>31</td>
<td>182</td>
<td>169</td>
</tr>
</tbody>
</table>

Three experiments were conducted to build the classification models for the five insurances. The first one used a single neural network with five outputs to classify the five insurances simultaneously. The second one used five neural networks (each with one output) to classify the five insurances independently. The first two experiments used the purchase records of primary and additional insurances as experimental data. The third experiment 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 (displayed at the end of this paper for a better article architecture).

**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. Each of the output nodes was associated with a single insurance.

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.

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

<table>
<thead>
<tr>
<th>Round</th>
<th>Integration of the five insurances</th>
<th>Individual Insurances</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Health</td>
<td>Accident</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Overall samples</td>
<td>61</td>
<td>85.6</td>
<td>97.6</td>
</tr>
<tr>
<td>2</td>
<td>Training (2/3)</td>
<td>64</td>
<td>96</td>
<td>95.8</td>
</tr>
<tr>
<td></td>
<td>Test (1/3)</td>
<td>26</td>
<td>75</td>
<td>86</td>
</tr>
<tr>
<td>3</td>
<td>Training (3/4)</td>
<td>77.3</td>
<td>90.2</td>
<td>95.5</td>
</tr>
<tr>
<td></td>
<td>Test (1/4)</td>
<td>33.3</td>
<td>74.6</td>
<td>84</td>
</tr>
</tbody>
</table>

5. Discussion

According to the experimental results shown in Table 2 and table 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 64.40%. 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 insurance and investment-oriented insurance.

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.

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.

Three experiments were conducted in this study. The first one used a single neural network with five output nodes to classify the five insurances simultaneously. The second one improves the classification accuracies of experiment 1 by using five neural networks (each with one output) to classify the five insurances independently. The first two experiments use the purchase records of primary and additional insurances as experimental data. The third experiment further improves the classification accuracies of experiment 2 by using the data of primary insurances purchases only.

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.

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. In addition, there are six input attributes in our experiments. We might consider removing some of attributes with less importance using some variable reduction methods. The factor analysis might be one of suitable techniques for doing that.

Reference


<table>
<thead>
<tr>
<th>Round</th>
<th>Integration of the five insurances (experiment 1)</th>
<th>Individual Insurances (experiment 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Overall samples</td>
<td>64.6</td>
</tr>
<tr>
<td>2</td>
<td>Training (2/3 of samples)</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>Test (1/3 of samples)</td>
<td>30</td>
</tr>
<tr>
<td>3</td>
<td>Training (3/4 of samples)</td>
<td>53</td>
</tr>
<tr>
<td></td>
<td>Test (1/4 of samples)</td>
<td>18.6</td>
</tr>
</tbody>
</table>