A TWO-STAGE DEEP NEURAL NETWORK TECHNIQUE FOR PREDICTING THE INSURANCE CROSS-SELLING OPPORTUNITY

Xiaoguang Tian
University of North Texas
1155 Union Circle #311160
Denton, TX 76203-5017
(972)559-9866
xiaoguang.tian@unt.edu

ABSTRACT

The benefits of insurance cross-selling can increase revenue and make the business profitable. In this study, I propose applying a two-stage deep neural networks approach to identify the customer who will purchase a new caravan insurance policy and improve the model performance through tuning the parameters. The results show that the approach is effective and outperforms baseline model on multiple indices.

INTRODUCTION

With the business revive and technology development, the insurance company CRM system currently stores large numbers of and complex customer data. It becomes a challenge for the company to find out the meaningful insights from the huge information. The application predictive modeling in CRM would help the company analyze and predict the customers' behaviors, and then acquire potential customer, retain existed customer, and add more value to the business.

Generally, an Artificial Neural Network (ANN) is based on a collection of many interconnected units called neurons (nodes). Each connection can transmit information results between neurons at different layers. The major computing components of an ANN include weights, summation transfer function, and activation function.

Compared to other learning techniques, neural networks can identify complicated nonlinear relationships and interactions between independent and dependent variables. They are effective to analyze noisy, incomplete, and less accurate data. Through tuning the parameters, such as number of hidden nodes in hidden layers, we can improve the performance of neural network.

The rest of the paper proceeds as follows. The second section provides a literature review on CRM, cross-selling, and ANN. I will discuss the data, proposed approach, results, and findings in the third section. The fourth section includes conclusions.

LITERATURE REVIEW

The previous insurance predictive modeling studies concentrated on customer segmentation, retention, profitability and satisfaction.

Florez and Ramon described a three-stage approach that combines marketing feature selection, customer segmentation through univariate and Oblique Decision Tree techniques, and a global cost–benefit function for measuring the success of the program. The result shows Decision Tree segmentation techniques achieve a higher performance and can easily be understood by managers (2009). One emerging area is the ability to dynamically segment users based on a machine learning technique called Latent Dirichlet Allocation (LDA). LDA might find a cluster of users that purchased certain brands of products, identify a particular user as belonging to that cluster (2015, Seth Earley).

The company can increase sales by focusing marketing resources on high quality customers. Smith et al. (2000) presented a case study involving understanding policyholders' retention pattern through various machine learning techniques, such as Logistic Regression, Neural network, and Decision Tree. From the knowledge of policyholders' possibility of canceling their policies, insurer can determine the cost of misclassification and the optimal pricing level, improving the market decisions.

Fang et al. (2016) applied random forecast regression, a method for big data analytics, to forecast insurance customer profitability. They found that customers' region, age, insurance status, sex, and customer source are the most important factors to predict insurance customer profitability according to the data they studied. Mehregan and Samizadeh (2012) applied K-means algorithm to identify the customers who purchase one or more insurance policies and find most applicable attributes which connect to the customers' purchase inclination.

Bockhorst et al. (2016) demonstrated a machine-learning-based framework to predict customer claim satisfaction by extracting relevant information from claim loss, notes and call and activity logs data. In order to extract information from unstructured text data, the claim handler notes were tokenized and processed by PoS tagger, generate the most frequent words ti-idf matrix approach, and reduce the dimensionality of attributes with Principal Component Analysis.

Most strategic decisions of cross-selling rely on managers' intuition and experience (Ansell, Harrison, & Archibald, 2007). In practice, one of the most important CRM decisions for the insurance company is figuring out who is the qualified customer to buy a specific product or service. The machine learning approach, especially the ANN, can be utilized to support the decision making process.

ANN is broadly applied in solving business problems. One particular type of ANN is called the Multilayer Perceptron (MLP). The most common method for training an MLP is through backpropagation, which computes the weights for a multilayer network and employs the gradient descent to look for the minimum squared error between the network output values and the target values for these outputs. Generally, the three major learning approaches for ANN are supervised learning, unsupervised learning, and reinforcement learning, which means ANN can be used across most of the tasks of machine learning.

Due to its powerful prediction ability and flexible application, ANN has been used in various fields of business and finance research to solve regression and classification problems, such as bankruptcy prediction, market share forecasting, stock performance, bond trade, and loan

application. Among these studies, the majority suggested that ANN performs as well or better than other machine learning/ statistical techniques (Wong & Selvi, 1998; Vellido, Lisboa, & Vaughan, 1999). However, neural networks have some disadvantages: the techniques are difficult to explain and interpret, the "black box" method limits people's understanding process and conclusion.

One of the most challenging task to fit an ANN is to determine the number of hidden layers and nodes in that hidden layer. We can pick the number of variables as the number of units in input layer: number of classes as the number of units in output layer. For hidden layers, the select strategy can be varied. Commonly, researchers use single layer neural networks to train the model (Sehgal et al., 2012; Zuriahati Mohd Yunos, Aida Ali, Siti Mariyam Shamsyuddin & Sallehuddin, 2016), applicable for easy task and reduce the training time but limiting the capability of the neural network.

In this study, I purposely apply a two stage deep neural networks approach to identify the customer who will purchase a new caravan insurance policy and improve the model performance through tuning the parameters.

EXPERIMENTS AND ANALYSIS

Dataset

This dataset I studied is from the CoIL 2000 Challenge and owned by the Dutch datamining company Sentient Machine Research. The dataset is based on real world business data and contains 86 variables which include the product purchase data and the socio-demographic data derived from zip area codes. The task is to identify who would be interested in purchasing a caravan policy. The names of products include private third party insurance, third party insurance (firms), car policy, and other 18 polices.

Experiment Settings

1. Sampling and partition original dataset to avoid imbalanced and overfitting problems.

2. In stage 1, I kept the input variables, standardization method and activation function in control for checking the outcomes caused by manipulating the number of layers and units. According to the different combination of layers and units (model parameters settings will be provided upon request) in each layers, I finally tested 7 models.

3. In stage 2, I applied ensemble approach to combine the two best models which are produced in stage 1 for optimizing the model performance.

Results and Analysis

In this study, networks were initially developed. From the modeling results of 7 networks, the best model is Model 2 and Model 1 which have the highest sensitivity and accuracy. The results are outlined in table (Table of results will be provided upon request) and will be discussed according to measurement indices.

CONCLUSION

In this study, ANN shows the powerful prediction ability to the insurance cross-selling problem. The technique can be widely used in solving various classification and prediction problems in the insurance domain.

Choosing the hidden layer is an effective strategy to enhance the performance of ANN models. From this study, deep neural network, which is ANN based on at least two hidden layers and large number of hidden units, we could gain more information from the insurance cross-selling data and identify more target consumers within the organization database. Ensemble model leverages the performance of the deep neural network and enables a more robust prediction and classification.

REFERENCES

References will be provided upon request.