Global Journal of Pure and Applied Mathematics.
ISSN 0973-1768 Volume 13, Number 6 (2017), pp. 1701-1710

http://www.ripublication.com

Artificial Intelligence for Estimation of Future Claim Frequency in Non-Life Insurance

K.M. Sakthivel and C.S. Rajitha

Department of Statistics, Bharathiar University
Coimbatore 641 046, Tamilnadu, India

Abstract

The essential feature of an insurance practice is to set the premium at the beginning of the insurance contract. To determine the correct premium for next year in an insurance company, precise and reliable estimate of the number of occurrence of claims and the total claim amounts is extremely important. Different methods are available in the literature for predicting the claim frequency of a policy for forthcoming years such as Generalized linear models (GLMs), Poisson regression models, Credibility models, Bayesian Models etc. But due to the heterogeneous nature of policies these methods does not produce exact and reliable prediction of future claim frequencies. Besides these conventional statistical methods depends largely on some limiting assumptions such as linearity, normality, independence among predictor variables and a pre-existing functional form relating the criterion variable and predictive variables etc. Recent studies in Artificial Intelligence show that Artificial Neural Networks (ANN) is powerful tools for prediction tasks due to their nonlinear nonparametric adaptive learning properties. In this paper, we developed the procedure for predicting the future claim frequency of insurance portfolio in general insurance using ANN with use of Bayesian credibility inputs with suitable illustration.

Keywords: Claim Frequency; Artificial Neural Networks; Bayesian Credibility; Artificial Intelligence.

INTRODUCTION

The fundamental role of insurance is to provide financial protection, offering a method of transferring the risk in exchange of an insurance premium. Therefore estimation of the right premium for the customers (policyholders) is one of the most important tasks in insurance business. Insurance companies provide insurance to the policy holders and in turn the policy holders have to pay insurance premium periodically. In a competitive market, charging a fair premium according to the expected loss of the policyholder is profitable for the insurer. The premium paid by the policyholders is determined from estimates of their expected claim frequencies and the severity of the claims in non-life insurance business. Therefore setting the precise and reliable estimates of the number of occurrence of claims has extreme importance. Different classical statistical methods have been applied in the industry for predicting the claim frequency of a policy for forthcoming years such as Poisson regression models, Generalized linear models (GLMs), Credibility models, Bayesian Models etc. Among these models Poisson model is considered as the most commonly used modeling representative of the frequency of claims in actuarial literature (Antonio and Valdez, 2012). In the case of non life insurance, According to Dionne and Vanasse (1989, 1992), Denuit and Lang (2004), Gourieroux and Jasiak (2004), Poisson regression model represents the main tool for the modeling claim counts (frequency) in non-life insurance. An important milestone in the development of Poisson regression models for claim counts is the contribution of Cameron and Trivedi (1998). They have managed to highlight the particularities of Poisson regression approach in modeling the claim frequency as a particular case of GLMs. However traditional methods used generalized linear models (Nelder and Wedderburn, 1972) for modeling the claim frequency (e.g. Renshaw, 1994; Haberman and Renshaw, 1996). To simultaneously model frequency and severity of insurance claims Jorgensen and de Souza (1994) and Smyth and Jorgensen (2002) used GLMs with a Tweedie distributed (Jorgensen, 1987, 1997) outcome. Yi Yang et.,al (2016) proposed a gradient tree-boosting algorithm and apply it to Tweedie compound Poisson models for predicting pure premiums.

For predicting future claim frequency of short term insurance contracts credibility theory is the most suitable technique. The technique calculates a claim frequency for a risk using two ingredients: past data from the risk itself and collateral data which allows an insurer to adjust future claim frequency based on past experience. The essential feature of a credibility claim frequency is that it is a linear function of the past data from the risk itself and that it allows for the claim frequency to be regularly updated as more data are collected in the future (Pacakova, 1997). Pacakova (2013) provides methods for permanently updated estimates of the number of claims and of the pure premium for each coming year in insurance company. But despite of the popularity of all these methods such as Poisson regression models, GLMs and Credibility models, these approaches rely on a set of stringent conditions such as linearity, independence, normality, etc. In reality, it may not possible to encounter ideal situation to meet and fulfill all these assumptions.

Neural networks (NNs) are being used as an alternative to all these traditional techniques and gaining popularity in recent years. Recent studies in Artificial Intelligence shows that ANN is powerful tools for prediction and classification, areas where regression models and other related statistical techniques have traditionally been used, due to their nonlinear nonparametric adaptive learning properties. NNs are data dependent and therefore, their performance improves with sample size. But statistical methods, such as Regression perform better only for extremely small sample size.

Although neural networks originated in mathematical neurobiology, the rather simplified practical models currently in use have moved steadily towards the field of statistics. The connection of neural networks to traditional statistical methods has been illustrated by many researchers. Cheng and Titterington (1994) discussed some thoughts on the future of the interface between neural networks and statistics and also made a detailed analysis and comparison of various neural network models with traditional statistical methods. They have shown strong connections of the feed forward neural networks with regression. Ripley (1994) classifies neural networks as one of a class of flexible nonlinear regression methods and also discusses the statistical aspects of neural networks. Sarle (1994) showed that the most commonly used artificial neural networks, called multilayer perceptrons, and are nothing more than nonlinear regression and discriminant models that can be implemented with standard statistical software. They have explained neural networks concepts and translate neural network jargon into statistical terminology and show the relationship between neural networks and statistical models.

In this paper, we have developed the procedure for predicting the future claim frequency of insurance portfolio in general insurance using ANN. Despite the apparent substantive and applied advantages of statistical models, Neural Networks methods have also gained popularity in recent years. These methods are particularly valuable when the functional relationship between independent and dependent variables are unknown and there are ample training and test data available for the process. NNs models also have high tolerance for noise in the data and complexity. Our research objective is to compare the predictive ability of NNs with Bayesian credibility inputs. In this paper, we used feed forward neural networks with the back-propagation algorithm to predict the claim frequency for forthcoming year using the Bayesian credibility inputs of the annual number of claims resulting from motor third-party liability insurance in an insurance company. This paper is organized as follows: In section 2 we provide a detailed description about NNs computation and various types of neural networks. Section 3 gives details of data and software used for this paper. Section 4 & 5 contains results and conclusions.

2. NEURAL NETWORKS

NNs have been shown to be very trusted systems in many applications such as prediction, classification, forecasting, and modeling due to their ability to learn from the data and also their nonparametric nature. The NNs solution may assist in predicting the reliable claim frequency for the forthcoming year.

2.1 Neural computation

Artificial neural network is a machine learning technique that is designed to model the way in which the brain performs a particular task or function of interest; the network is usually implemented by using electronic components or is simulated in software on a digital computer (Simon Haykin, 2001). An ANN, usually configured for a specific application, is composed of a large number of highly interconnected parallel processing elements, called neurons, working in agreement to solve specific problems, through a learning process (Ibiwoye et.al., 2012). An artificial neuron is an eclectic simulation of biological neuron, and it consists of its own dendrites, synapses, cell body and axon terminals. It receives stimulation from nearby cells, or from its environment, and generates a modified action potential or nerve signal (Ajibola et al., 2011). According to Stergious and Siganos (2007), ANN approach has a unique capability for deriving meaning from complicated or imprecise data and is useful in detecting patterns or trends that are too complex to be noticed by humans or other computer techniques. Agatonovic-Kustrins et al. (2000) explained that the behavior of a neural network is determined by the transfer functions of its neurons, by the learning rule, and by the architecture itself. Figure1 is the architectural framework of the commonly used artificial neural network consisting of layers of input units connected to layer of hidden units which are connected to a layer of output units.

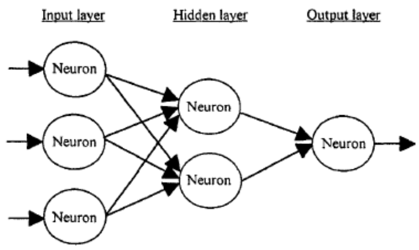

Figure 1

The behavior of an ANN depends on both the weights and the input-output (transfer function) that is specified for the units (Stergious and Siganos, 2007).These functions fall into one of three categories namely: linear, threshold and sigmoid functions. Several possibilities of using transfer functions of different types in neural models are discussed in (Duch and Jankwoski, 2001). During training, the inter-unit connections are optimized until the error in predictions is minimized and the network reaches the specified level of accuracy. Nguyen, (2005) investigated the predictive power of various neural network models in predicting corporate failure.

NNs are mostly good at recognizing complex patterns. The central idea of a NNs is to extract linear combinations of the inputs as derived features, and then model the target variable as a nonlinear function of these features (Hastie, Tibshirani, and Friedman

2004). The NNs algorithm family is quite large. NNs are a statistical model that can be represented by a network diagram.

2.2 Neural network types

In the last few decades different types of neural networks have been investigated. Among these the feed forward NNs was the simplest type of artificial NNs. In the feed forward network, the network moves in the forward direction. In the case of feedback networks the networks include connections back to previous layers or even back to the neuron itself. The former class of the network has been used in this paper.

3. DATA

In the present paper, we compare the predicted claim frequency using two methods, one is Bayesian credibility model (Poisson/gamma model) and another one is ANN. The data set consists of the annual number of claims resulting from motor third party liability insurance in insurance company in the years 2006-2011 (Pacakova ,2013).

3.1 Bayesian credibility model

In the Bayesian credibility model for calculating the claim frequencies use two ingredients, past data from the risk itself and the collateral data. The credibility formula for estimation of claim frequency in next year is:

$$P_C = Z\,\bar{x} + (1-Z)\,\mu$$

where $\bar{x}$ is estimation based on past data from the own data in insurance company, or risk, and μ is estimation based on collateral data and Z is a number between zero and one, known as the credibility factor. Credibility factor Z is a measure of how much reliance the company is prepared to place on the data from the policy itself. The claim frequency rate for a class of insurance business may lie anywhere between 0 and $+\infty$. An insurer with a large experience may quite accurately estimate the rate.

Here we assume the distribution of a number of claims is the Poisson distribution with a fixed but unknown parameter λ, so the likelihood function has the form

$$f(x/\lambda) = \prod_{i=1}^{n} \frac{\lambda^{x_i}}{x_i!} e^{-\lambda} = c_1 e^{-\lambda n} \lambda^{\sum_{i=1}^{n} x_i} ; x = 0,1,2,\ldots; \quad \lambda > 0$$

Our objective is to estimate the unknown parameter λ. The gamma distribution (prior) may be convenient for representing uncertainty in a current estimate of the claim frequency rate. This distribution is over the whole positive range from 0 to $+\infty$, and the mean α/β can be set equal to the current best estimate. Uncertainty is represented by variance α/β^2 of the gamma distribution $G(\alpha;\beta)$.

In Poisson /gamma model for claim numbers, we have assumed prior knowledge about the unknown parameter λ summarized by gamma distribution $G(\alpha,\beta)$ with known parameters α and β. By assumption the density function of the prior $G(\alpha;\beta)$ distribution is

$$f(\lambda)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}\lambda^{\alpha-1}e^{-\lambda\beta}=c_2e^{-\lambda\beta}\lambda^{\alpha-1}$$

By Bayes' theorem, we get the posterior density of λ, given **x = (** $x_1,x_2,\ldots x_n$ **)**, in the form

$$f(\lambda/x)\propto e^{-\lambda n}\lambda^{\sum x_i}\lambda^{-\alpha-1}e^{-\lambda\beta}=e^{-\lambda(\beta+n)}\lambda^{\alpha+\sum x_i-1}$$

that is the gamma distribution with new parameters

$$\alpha_1=\alpha+\sum x_i$$

$$\beta_1=\beta+n$$

The Bayesian estimate of λ, with respect to a quadratic loss function, given these data, is the mean of the posterior density of λ.

$$\lambda_B=E(\lambda/x)$$

Thus the Bayesian estimator of λ using the quadratic loss is

$$\lambda_B=\frac{\alpha+\sum_{i=1}^{n}x_i}{\beta+n}=\frac{\alpha+n\bar{x}}{\beta+n}$$

If we put factor credibility $Z=n/(\beta+n)$. Then $\lambda_B=E(\lambda/x)=Z\ \bar{x}+(1-Z)\,\mu$ which is the credibility formula for updating claim frequency rates.

It can be seen from the credibility factor expression, since n is non-negative and β is positive, that Z is in the range zero to one and it is increasing function of n. If no past data from the risk itself are available, then the best estimate of λ is α/β, the mean of the prior gamma distribution. It can be seen that Z does not take the value one for any finite value of n. The value of Z depends on the amount of data available for the risk n, and the collateral information through β, which reflect the variance α/β^2 of the prior distribution.

In Poisson/gamma model for claim numbers, we have assumed prior knowledge about the unknown parameter (annual claim rate) λ summarized by gamma distribution $G(\alpha,\beta)$ with known parameters α = 8400 and β = 0.4.

3.2 Claim frequency rate using neural network

Here we predict the claim frequency of forthcoming year using ANN with the Bayesian credibility inputs of the Poisson/gamma model (Pacakova, 2013). The Bayesian credibility inputs contain the annual number of claims(actual claim frequency) resulting from motor third party liability insurance in the years 2006-2011 and the estimated credibility factor and estimated annual claim numbers calculated using the Bayesian credibility method. This data set is considered as test set (actual number of claims and credibility factor is considered as inputs and the estimated number of claims as desired output). The train set is generated by using the interpolation. The generated train set consists of 100 observations for each variable, and are trained using the neuralnet package in R software with feed forward NN with back propagation algorithm. During the training phase the model learns how to use some of the fields in a record to predict the value of another field. Once it is trained and tested, it can be given new input information to predict an output.

4. RESULTS

Table 1 presents the prediction of claim frequency using Bayesian credibility method and ANN using claim number as covariate. In table 1 the column denoted as λ_B contains values of Bayes' estimators of annual claim frequency rates x_i for each year i based on $(i-1)$ past observations by using the credibility formula for updating claim frequency rates. And z_i denote the credibility factor. The column denoted as λ_{ANN} contains the values of annual claim frequency rates calculated using ANN model with a single hidden layer having 15 neurons.

Table 1. Prediction of number of claims using Bayesian credibility and ANN

year	x_i	z_i	λ_B	λ_{ANN}
2006	24954	0.6	21000	22954
2007	23166	0.71429	23824	21397
2008	19402	0.83333	23550	21089
2009	18658	0.88235	22330	22362
2010	19142	0.90909	21495	21019
2011	20618	0.92593	21060	23675
2012	-	0.93750	20991	20592

The purpose of network training is to estimate the weight matrices such that an overall error measure such as the mean squared errors (MSE) or sum of squared errors (SSE) is minimized. MSE can be defined as

$$MSE = \frac{1}{N}\sum_{j=1}^{N}(a_j - y_j)^2$$

a_j and y_j represent the target value and network output for the jth training pattern respectively, and N is the number of training patterns.

Table 2. Standardized MSE of Bayesian credibility method and ANN

Method of estimating claim frequency	**MSE**
Bayesian Credibility	1.911
Artificial Neural Network	1.195

Table 2 shows the MSE of actual values (x_i) of annual claim frequency rate with predicted values of annual claim frequency rate using Bayesian credibility method (λ_B) and ANN (λ_{ANN}) with covariate as actual claim frequency. And also shows the MSE of predicted future claim frequency between outputs of Bayesian credibility and ANN. The results shows that MSE of ANN estimation procedure is small when compared to that of Bayesian credibility.

5. CONCLUSION

As we know the prediction of the future claim frequency is substantial in insurance practice to set the premium at the beginning of the insurance contract. Though various methods have been developed for different circumstances for prediction purposes, that methods depends on some limiting assumptions such as linearity, normality, independence etc and also takes comparatively more time than rather simplified practical ANN model for prediction. Since NNs are trained using the already available original data the error may be reduced and also takes less time compared to other methods for prediction purposes. We compared the ANN with Bayesian credibility method, a statistical method for predicting the future claim frequency. The results confirmed that ANN provides significantly better estimate of the actual claim frequency compared to Bayesian credibility in case of Poisson/gamma model for non-life insurance.

REFERENCE

[1] Agatonovic Kustrins, S., and Beresford, R., 2000, "Basic Concepts of Artificial Neural Network (ANN) Modeling and its Application in Pharmaceutical Research," Journal of Pharmaceutical Biomedical Analysis., 22 (5), pp. 717-27.

[2] Ajibola, O. O. E., Olunloyo, V. O. S., and Ibidapo- Obe, O., 2011, "Artificial Neural Network Simulation of Arm Gait in Huntington's Disease Patient," International Journal of Biomechatronics and Biomedical Robotics., 1 (3), pp. 133-140.

[3] Antonio, K., and Valdez, E. A., 2012, "Statistical Concepts of a Priori and a Posteriori Risk Classification in Insurance," Advances in Statistical Analysis., 96(2), pp. 187-224.

[4] Cameron, A. C., and Trivedi, P. K., 1998, Regression Analysis of Count Data, Cambridge University Press, Cambridge, UK.

[5] Cheng, B., and Titterington, D. M., 1994, "Neural networks: A review from a statistical perspective," Statistical Science., 9(1), pp. 2–30.

[6] Denuit, M., and Lang, S., 2004, "Nonlife Ratemaking with Bayesian GAM's," Insurance: Mathematics and Economics., 35(3), pp. 627-647.

[7] Dionne, G., and Vanasse, C., 1989, "A Generalization of Automobile Insurance Rating Models: the Negative Binomial Distribution with a Regression Component," ASTIN Bulletin., 19(2), pp. 199-212.

[9] Dionne, G., and Vanasse, C., 1992, "Automobile Insurance Ratemaking in the Presence of Asymmetrical Information," Journal of Applied Econometrics., 7(2), pp. 149-165.

[10] Duch, N., and Jankowski, N., 2001, "Transfer Functions: Hidden Possibilities for Better Neutral Networks," Proc. 9th European Symposium on Artificial neutral Networks, Belgium, pp. 81-94.

[11] Gourieroux, C., and Jasiak, J., 2004, "Heterogeneous Model with Application to Car Insurance," Insurance: Mathematics and Economics., 34(2), pp. 177-192.

[12] Haberman, S., and Renshaw, A. E., 1996, "Generalized linear models and actuarial science," Statistician., 45(4), pp. 407-436.

[13] Hastie, T. J., and Tibshirani, R. J., 1990, Generalized additive models, CRC Press, vol. 43.

[14] Jorgensen, B., 1987, "Exponential dispersion models,". Journal of the Royal Statistical Society. Series B (Methodological)., 49(2), pp. 127-162.

[15] Jorgensen, B., 1997, The theory of dispersion models, CRC Press, vol. 76

[16] Ibiwoye, A., Ajibola, O. O. E., and Sogunro, A. B., 2012, "Artificial Neural Network Model for Predicting Insurance Insolvency," International Journal of Management and Business. Res., 2 (1), pp. 59- 68.

[17] Jorgensen, B., and De Souza, M. C., 1994, "Fitting Tweedie's compound Poisson model to insurance claims data," Scandinavian Actuarial Journal., pp. 69-93.

[18] Nelder, J., and Wedderburn, R., 1972, "Generalized Linear Models," Journal of the Royal Statistical Society. Series A (General)., 135(3), pp. 370-384.

[19] Nguyen, H.G., 2005, "Using Neutral Work in Predicting Corporate Failure," Journal of Social Sciences., 1(4), pp. 199-202.

[20] Ohlsson, E., and Johansson, B., 2010, Non-life insurance pricing with generalized linear models, Springer.

[21] Pacakova, V., 1997, "The Bayesian Inference in Actuarial Sciences," Central European Journal for Operations Research and Economics., 5, (3-4), pp. 255-268

[22] Pacakova,V., 2013, "Credibility models for permanently updated estimates in insurance," International Journal of Mathematical Models and Methods in Applied Sciences., 7, pp. 333- 340.

[23] Renshaw, A. E., 1994, "Modelling the claims process in the presence of covariates," ASTIN Bulletin., 24(2), pp. 265-285.

[24] Ripley, B. D., 1994, "Neural networks and related methods for classification," Journal of the Royal Statistical Society. Series B (Methodological)., 56(3), pp. 409–456.

[25] Sarle, W. S., 1994, "Neural networks and statistical models," Proc.19th annual SAS users group international conference.

[26] Simon Haykin., 2001, Neural Networks, A comprehensive foundation, Second Edition, Prentice Hall, Pearson Education, Singapore.

[27] Smyth, G., and Jorgensen, B., 2002, "Fitting Tweedie's Compound Poisson Model to Insurance Claims Data: Dispersion Modelling," ASTIN Bulletin., 32(1), pp. 143-157.

[28] Stergious, C., and Siganos, D., 2007, "Neural Networks," Available: www.Doc.k.ac.uk/and/surprise96/journal /Vol4/cs11/report.html

[29] Yi Yang.,Wei Qian., and Hui Zou., 2016, "Insurance Premium Prediction via Gradient Tree-Boosted Tweedie Compound Poisson Models," Journal of Business and Economic Statistics, 34(3), pp. 1-45.