International Journal of Integrated Engineering:
Special Issue 2018: Data Information Engineering, Vol. 10 No. 6 (2018) p. 193-196.
 DOI: https://doi.org/10.30880/ijie.2018.10.06.027

Comparison between ANN and Multiple Linear Regression Models for Prediction of Warranty Cost

Mohd Faaizie Darmawan[1*], NurIzzati Jamahir[1], Rd. Rohmat Saedudin[2], Shahreen Kasim[3]

[1]Soft Computing & Intelligent System (SPINT), Faculty of Computer Systems & Software Engineering, University Malaysia Pahang, Lebuh raya Tun Razak, Gambang, 26300, Kuantan, Pahang, Malaysia
[2]School of Industrial Engineering, Telkom University, 40257 Bandung, West Java, Indonesia
[3]Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Batu Pahat, Johor, Malaysia.

Received 28 June 2018; accepted 5 August 2018, available online 24 August 2018

Abstract: Nowadays, warranty has its own priority in business strategy for a manufacturer to protect their benefit as well as the intense competition between the manufacturers. In fact, warranty is a contract between manufacturer and buyer in which the manufacturer gives the buyer certain assurances as the condition of the property being sold. In industry, an accurate prediction of warranty costs is often counted by the manufacturer. Underestimation or overestimation of the warranty cost may have a high influence to the manufacturers. This paper presents a methodology to adapt historical maintenance warranty data with comparison between Artificial Neural Network (ANN) and multiple linear regression approach.

Keywords: Artificial Neural Network, Warranty Cost, Computational Intelligent

1. Introduction

Warranty has its own priority in business strategy for a manufacturer to protect their benefit as well as the intense competition between the manufacturers. In fact, warranty is a contract between manufacturer and buyer in which the manufacturer gives the buyer certain assurances as the condition of the property being sold. In industry, an accurate prediction of optimal warranty period and warranty costs is often counted by the manufacturer. A warranty period may be unprofitable for the manufacturers if the choice of duration given is either too short or too long. Same thing goes to warranty cost which is an underestimation or overestimation of the warranty cost may have a high influence to the manufacturers.

The great market competition makes that the companies look for high quality of products. Basically, there are six types of warranty that can be offered either by manufacturer or dealers of a product or services which are basic warranty, extended warranty, warranty for used, repair limit warranty, service warranty and lifetime warranty. The different types of warranty policy were established in order to fulfill the demand of manufacturers and the requirement of buyers so that a win-win situation could be acquired. This study will apply basic warranty which mean that it will covers all factory-installed parts against any type of defects in their manufacturing and workmanship.

The area of warranty has been studied by researchers from many different areas such as economic, engineering, and statistics. A number of models have been used as a model in solving warranty problem. In past few years, there has been an increased used of statistical models instead of soft computing models in warranty related applications. However, soft computing models have been used by many researchers in the other research area which can provide some feasible solutions for the complex real-world problems. For warranty problem, there are several studies in warranty problem specifically by using soft computing model [1, 7, 8,9, 10, 13, 14].

In this paper, we proposed ANN and Multiple linear regression model for prediction of warranty cost. The study is organized as follows. Section 2 gives a brief literature review on models used in warranty problem. Section 3 describes data collection and proposed models for this research. Section 4 discusses results of both proposed models. Section 5 is discussion while Section 6 for conclusions.

2. Literature Review

Reference [2] noted that until 1940's, warranty studies were carried out by researcher in the legal field and therefore its appeared only in journals of law. Basically, warranty define as a contractual agreement that requires the manufacturer to rectify a failed item, either through repair or replacement, failure should not be attribute to reckless use, it must occur within the period prescribed by warranty [3].

*Correspondingauthor: faaizie@ump.edu.my

A vast review on warranty can be found. It is stated by [4] about 1500 articles in this area are listed. For example, [5] proposed a mixture model for automobile warranty data. This model fits warranty claims as a mixture of manufacturing or assembly defects (quality problems) and usage related failures (reliability problems). Reference [6] proposed modelling and optimization of extended warranties using probabilistic design.

A number of models in solving warranty problem have been increasing but applications in soft computing in warranty area receive less intention. There are several recent soft computing models applied in warranty problem. For example, [7] suggested different models which are suitable for modelling a two-dimensional warranty plan, and chooses a suitable fuzzy model to handle vague data. Reference [8] studied on application of fuzzy feedback control for warranty claim. They proposed a model that captures fuzzy events to determine the optimal detection of warranty claims data. In addition, neural network has been applied by [9] and [10] in order to predict or forecast automobile warranty claim data.

3. Materials and Methods

In this study, the historical maintenance data is drawn from automobile Malaysia company. The data used is single product which is MTB 150dx. There are about 726 historical data recorded from the 102 samples vehicle beginning from the first inspection to the end inspection for each vehicle. Some element that can be found in this historical data are the age, vehicle mileage, number of failure and defect, and cost of every inspection. The age, vehicle mileage, number of failure, and number of defect is considered as independent variable while the cost of inspection is considered as dependent variable.

3.1 Multiple Linear Regression

Multiple regression, a time-honoured mathematical model going back to Pearson's 1908 use of it, is employed to predict the variance in an interval dependency, based on linear combinations of interval, dichotomous, or dummy independent variables.

Simple regression analysis was performed at the first stage of the analysis which is to learn more about the relationship between several independent or predictive variables and a dependent or criterion variable. The relationship is measured by correlation coefficient (R2) which is an indicator of correlation strength. After that, the mean square error is calculated for each combination of variables used as shown in Table 1. The best combination of variable is then chosen based on the highest R2 value and the lowest MSE value. This combination is then used by the ANN model for the same purpose, predict the warranty cost. Accuracy of the multiple linear regression model is then calculated based on the combination of variables chosen using Equation 1, and will be compared with the ANN model for comparison purpose.

$$Accuracy = 100 \times \left[1 - \frac{|z_f - z_e|}{|z_e|}\right] \quad (1)$$

where z_f is derived output and z_e is expected output.

3.2 ANN Model

ANN are inspired by biological role of nerve cells in the human brain. ANN consist of a group of a number of interconnected cells called as neurons with weight running together to create learning in machines [11]. Basically, ANN consist of three interconnected layers which are input, hidden and output layer. The input and output layer consist of a collection of neurons which represents input and output variables or data. The hidden layer also consists of a series of specific neurons and is connected in between the input and output layer. There are several studies in neuro of ANN in this area such as in [12].

For this case study, the best combination variables selected for multiple linear regression is chosen as input for ANN model. Four different models are used in predicting the warranty cost based on popular back-propagation training algorithms which are Lavenberg-Marquardt back propagation algorithm (*trainlm*), Gradient descent back propagation algorithm (*traingd*), Gradient descent with momentum back propagation algorithm (*traingdm*) and Gradient descent with momentum and adaptive learning back propagation algorithm (*traingdx*). According to [13] the best combination transfer function is *Tansig-Logsig-Tansig*with three hidden layers. Therefore, this combination transfer function was used and tested with difference learning algorithm. To select the best model, each model will produce the MSE value for training and testing dataset, and the model that produce the lowest MSE value for both dataset is then chosen as the best model for the ANN model in predicting the warranty cost. The accuracy for the chosen model is then calculated using the same Equation 1, and will be compared with the multiple linear regression for comparison purpose.

4. Summary

Table 1 shows the result of relationship using multiple linear regression between the predictive variable which is age, mileage, defect and failure, and dependent variable which is cost of inspection. From the table, the combination of all variables produced the lowest MSE value and the highest R2 value. Thus, this combination of variables is then chosen to calculate accuracy of prediction for warranty cost using multiple linear regression model. The accuracy is then tabulated in Table 3 for comparison purpose with the ANN model.

Table 1: Variable Selection using Multiple Linear Regression

Age	Mileage	Defect	Failure	Regression	
				MSE	(R^2)
✓	✓	✓	✓	**0.092**	**0.698**
✓	✓	✓		0.112	0.554
✓		✓	✓	0.093	0.692

✓	✓		✓	0.093	0.691
	✓	✓	✓	0.092	0.697
✓	✓			0.114	0.539
✓		✓		0.129	0.413
✓			✓	0.096	0.674
	✓	✓		0.113	0.550
	✓		✓	0.094	0.690
		✓	✓	0.094	0.686

Table 2 shows results obtained from the four ANN models chosen based on the learning algorithm, and the MSE value for both training and testing dataset. From the table, its clearly shows that *trainlm* is the best learning algorithm in this study with gives the lowest MSE value for both datasets. Thus, the *trainlm* model is chosen as the best model for ANN model, and the calculated accuracy for this model is tabulated in the same Table 3 for comparison purpose.

Table 2: Results obtained by the four ANN models and their performance evaluation

Learning Algorithm	MSE (training)	MSE (testing)
Trainlm	**0.037**	**0.0424**
Traingd	0.0437	0.0473
Traigdm	0.0396	0.0467
Traingdx	0.0432	0.051

Table 3 shows summary of the both multiple linear regression and ANN model based on performance measurement, accuracy, in predicting warranty cost. The table clearly shows that the predictive performance of ANN model is obviously higher than multiple linear regressions. The results of analysis show that ANN model has the best efficiency in comparison the multiple linear regressions.

Table 3: Accuracy produced by both models based on testing dataset, in predicting warranty cost

Models	Accuracy (%)
ANN	96.30
Multiple Regression	90.73

5. Discussion

ANN and multiple regressions are used in this study as the borderline or benchmark in investigating the prediction of data by comparing the soft computing modelwith one mathematical model. A number of researchers have employed ANN model as soft computing model in predicting warranty cost such as by [13] and [14].

The purpose of this analysis is to assess the performance and the accuracy of the multiple linear regression andANNmodel in dealing with warranty problem. Based on the results produced by multiple linear regression, the combination of all independent variables produced the best results for both MSE and R2. This combination is then used for ANN model using four training algorithms, to see the performance of each algorithm in predicting the warranty cost. Based on the results produced by ANN model, the *trainlm* algorithm is the best algorithm in predicting the warranty cost.

For overall, the performance measurement produced by the multiple linear regression and ANN model can be used to predict warranty cost problem with high accuracy. However, for comparison purpose, the ANN produced better performance than the multiple linear regression with 96.30 percent of accuracy performance.

6. Conclusion

In this study, ANN and multiple linear regression model were utilized to predict warranty cost where four effective parameters were considered as input to the both models. For the dataset used, 726 samples of data were used from the historical warranty data. According to the performance of both models, the prediction performance of ANN model is better than the multiple linear regression models. Thus, the ANN model is preferred to be used to predict warranty cost compared to multiple linear regression. However, in future development, this research can be improved by studying various other case studies available such as [15]-[20].

Acknowledgement

This study thanked to UMP RDU under grant number RDU1703127, for their support.

References

[1] Majid H.A., Kasim N.H., Jamahir N.I., Samah A.A. (2012). SoftComputingMethods in WarrantyProblems: ReviewandRecentApplications (2003-2012). *InternationalJournalofComputerScience Issue.*190-196.

[2] Murthy D.N.P, Blischke W.R. WarrantyManagementandProductManufacture, Springer, 2006.

[3] Singpurwalla N.D. and Wilson S.P. (1994). Software reliability modeling. International Statistical Review.(62):289.

[4] Murthy D.N.P and Djamaludin I. (2002). New product warranty: A literaturwe review. 231-260.

[5] MajeskeK.D (2003). A mixture model for automobile warranty data. *Reliability Engineering and System Safety.71-77.*

[6] Vittal S. and Philips R. (2007),Modeling and Optimization of Extended Warranties Using Probabilistic Design.*IEEE.*41-47.

[7] Lee S.H., Lee J.H., Park S.B., Lee M.T., Lee S.J., Kim B.K.(2008), A Fuzzy Reasoning Model of Two-dimensional Warranty System. *International Conference on Advanced Language Processing and Web Information Technology*.287-292.

[8] Lee S. H., Lee S. J., and Moon K. L. (2011). Application of Fuzzy Feedback Control for Warranty Claim. *SCI*, 279-288.

[9] Lee S. H., Seo, S. C., Yeom, S. J., Moon, K. I., Kang, M. S. and Kim, B. G. (2007). A Study on Warning/ Detection Degree of Warranty Claims Data using Neural Network Learning. *Proceeding of Sixth International Conference on Advanced Language Processing and Web Information Technology (ALPIT)*, 492-497.

[10]Rai, B. K. and Singh, N. (2009). Reliability Analysis and Prediction with Warranty Data. Boca Raton. Taylor and Francis Group, LLC.

[11]Karray, F. O., andSilva, C. D. (2004). SoftComputingandIntelligentSystemsDesign. London. PearsonEducationLimited.

[12]Hrycej T. andGrabert M. (2007), WarrantyCostForecastBasedon Car FailureData. InternationalJointConferenceonNeuralNetworks.

[13] Majid, H. A., Ang, J. C., and Samah, A. A (2011). "Modeling of Two-dimensional Warranty Policy using Artificial Neural Network (ANN) Approach," Journal of Computing, 3(12), 48 – 57

[14] Stamenkovic, D. D., and Popovic, V. M. (2013). "Warranty optimisation based on the prediction of costs to the manufacturer using neural network model and Monte Carlo simulation," International Journal of Systems Science, 46 (3), 535-545

[15] Darmawan, M. F., Hasan, H., Sadimon, S., Yusuf, S. M., and Haron, H. (2018). "A Hybrid Artificial Intelligent System for Age Estimation Based on Length of Left Hand Bone," Advanced Science Letters, 24(2), 1047-1051

[16] Ismail, M.A., Mezhuyey, V., Moorthy, K., Kasim, S., and Ibrahim, A. O. (2017). "Optimisation of Biochemical Systems Production using Hybrid of Newton Method, Differential Evolution Algorithm and Cooperative Coevolution Algorithm", Indonesian Journal of Electrical Engineering and Computer Science, 8, 27-35

[17] Darmawan, M. F., Yusuf, S. M., Rozi, M. A., and Haron, H. (2015). "Hybrid PSO-ANN for sex estimation based on length of left hand bone," IEEE Student Conference on Research and Development (SCOReD), 478-483

[18] Ismail, M. A., Mezhuyey, V., Deris, S., Mohamad, M. S., Kasim, S., and Saedudin, R. R. (2017). "Multi-objective Optimization of Biochemical System Production Using an Improve Newton Competitive Differential Evolution Method", International Journal on Advanced Science, Engineering and Information Technology, 7, 1535-1542

[19] Ismail, M. A., Deris, S., Mohamad, M. S., Isa, M. A., Abdullah, A., Remli, M. A., and Mohi-Aldeen, S. M. (2015). "A Hybrid of Optimization Method for Multi-Objective Constraint Optimization of Biochemical System Production", Journal of Theoretical and Applied Information Technology, 81, 502-513

[20] A. Adib, A. Banetamem, A. Navaseri, (2017), "Comparison between Results of Different Methods of Determination of Water Surface Elevation in Tidal Rivers and Determination of the Best Method", International Journal of Integrated Engineering,9,1-9.