© 2017 IJSRSET | Volume 3 | Issue 5 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
Themed Section: Engineering and Technology

Study of Road Accident Prediction Model at Accident Blackspot Area: A Case Study at Selangor

Haikal Aiman Hartika*[1], Mohd Zakwan Ramli[1], Muhamad Zaihafiz Zainal Abidin[2], Mohd Hafiz Zawawi[1]

*[1]Civil Engineering Department, College of Engineering, Universiti Tenaga Nasional, Kajang, Selangor, Malaysia
[2]Faculty of Architecture and Built Environment, Infrastructure University Of Kuala Lumpur, Kajang, Selangor, Malaysia

ABSTRACT

This study is focusing in the accident prediction model by using Multiple Linear Regression and Artificial Neural Network, as well as investigating the effectiveness of both predictive analyses. To meet these objectives, the historical data of accident was analysed with the on-site data for the purpose of relating them with the accident occurrence. Also, the accident historical data and the on-site data were analysed using two predictive models to predict the number of accident in current year. After the analysis, it is revealed that the outcomes obtained from both models were different with greater predicted value from the regression analysis due to greater errors being produced. Less predicted value, greater correlation and less error have been shown by the Artificial Neural Network analysis, concluding that it is more suitable for accident prediction.

Keywords: ANN, MLR, prediction model, and road accident.

I. INTRODUCTION

A preliminary hypothesis found that the increasing number of vehicles in the country contributes to road accident. Every year, there are an increasing number of vehicles on the road thus the road cannot accommodate such high traffic volume and this causes congestion. From 2010 to 2015, there is an increment number of drivers, according to a statistic by Road Transport Department Malaysia (JPJ) [1]. This is not including those who are driving without driving license and those with foreign driving license. Besides, the migration of people to the urban area also contributes to the problem as the migration increases the number of vehicles on the urban road plus they are unfamiliar with the city area and the roads in the city. Other than that, improper town planning also makes the problem worse. The engineers involved in the past projects did not have much long term visions on the planning especially in designing the road.

Transportation is a fundamental part of modern existence, linking the various activities in which people participate especially at home, work, school and go to shopping or recreation [2]. This is further supported by that transportation is vital both to the economic success and to the quality of life in urban and rural areas [3]. In addition to that, the main function of transportation system is to provide effective and secure movement of goods and passenger from one place to another so that economic development can steadily improving since it is directly and strongly related to the availability of transportation [4].

Although the transportation efficiency has improves the overall transportation system, there exists some issues and challenges which brought some adverse effects. This is due to the accelerated increases in the practice of own transport in Malaysia stimulated by inadequate public transport produced some issues such as traffic congestion, accidents, inadequate parking space and air pollution [5]. However, traffic congestion and road traffic accident are the two main issues that came into highlights today as they are interrelated to each other.

Road accident is the main concern regarding the issue of road safety. It is defined as the unexpected and unplanned incident involving collision of one or more parties which includes vehicles and pedestrian. This is previously explained by [6], whom defined road accident as a collision between several vehicles or at

least a moving vehicle on public or private road whereby damage or injury is affecting any person, property, vehicle, structure or animal and is recorded by the police due the negligence by any party or due to environmental factor. Road accident has caused many prominent problems. One of those is the economic losses. This is proven in China, where accidents in the highways cause huge economic losses to the country due to increasing number of accidents over time [7].

Other than that, road accidents are also one of the major causes of the unnatural loses of human beings in the world [8]. The rates of road accidents are increasing significantly in developing countries due to the increasing number of vehicles and inefficient drivers on the road, as well as the poor conditions and maintenance of the roads [9]. However, the crisis is not faced by most of the developed countries. This can be shown in Sudan, where the accident casualties are among the dominant causes of death in the age group of 20 to 60 with 61% casualties. The fatality rate of 35 per 10,000 vehicles is among the highest in the world despite low car ownership of a vehicle per one hundred person [10]. Also, traffic accidents are one of the main external causes of depression and fatality in Brazil, a country that has experienced rapid growth in the number of cars and motorcycles in the last five decades [11]. Furthermore, in Malaysia, road accidents are one of the major contributors of human life loss as there were 341,252 accidents recorded in 2006, resulting in an average 18 deaths from road accidents every single day [2]. However, the preliminary hypothesis is just a general assumption on why the accident problem arises. Therefore, it is important to study on the accident trends and predict the factors involved so that preventive measures can be taken in the future to save more lives from being killed by the road accidents. For this purpose, accident records from previous years are essential to investigate the trend as well as come out with the improvement. Also, the approach of predicting the accident numbers in the future might assist the responsible authorities to mitigate the problem. To make the predicting approach to be success, accident prediction models are preferred rather than using conventional method to analyse the accident trends and factors. This will analyse more accurate data on what causes the accident and how to reduce the problem in order to save lives.

In order to make a more accurate prediction, the predictive analysis shall be used in an exact road. For instance, every road which has the accident prone area shall use different predictive analysis to make the prediction more accurate. This means that the predictive analysis apply to a specific road, not in a general manner. Generally, there are many models available such as the Artificial Neural Network (ANN), Multiple Linear Regression (MLR), support vector and other methods. Each method has their own approach which produced different advantages and disadvantages according to situation given.

Efficiency of Prediction Model

Multiple linear regressions (MLR) are the type of regression containing two variables, which are dependent and independent respectively. It is an extent of simple linear regression. It is used for prediction of a variable based on the value of two or more variables. The variable produced is referred to as dependent variable, outcome, target or criterion variable. The variables that were used for prediction are called the independent variables, predictor, explanatory or regress or variables. In this study, the variable produced was the predicted number of accidents based on several variables such as the vehicular speed, gap between vehicles and traffic volume. Multiple Linear Regression (MLR) has been the most popular technique in developing injury severity prediction models [2]. This is further clarified by [12], whom stated that the dependent variables increase when the degree of independent variables increase.

Recently, the developments of another predictive modelling called Artificial Neural Networks (ANN) have gained more attentions from researchers across the globe. It is a computer model designed for knowledge processing and prediction. ANN has been applied to establish accident models in road safety studies. ANN models was developed to predict driver injury severity in traffic accidents at signalized intersections [13]. Other than that, ANN model were developed by applied historical data between 1986 and 2005 at 3 metropolitan cities of Turkey for estimation number of death due to road accident [14].

Therefore, this study is conducted by applying both models of Multiple Linear Regression (MLR) and Artificial Neural Network (ANN) in order to determine

the effectiveness of the modelling thus finding the most suitable model to be applied in this study.

In order to determine the most suitable model for this study, the comparison in terms of correlation coefficient and mean square error between the two predictive models will be conducted, in which high value of correlation and low value of mean square error would be preferred as the most suitable model. These two elements were chosen to determine the accuracy because they were adopted as criterion in for such analysis [15].

II. METHODS AND MATERIAL

The study was conducted along Persiaran Jubli Perak, Shah Alam, Selangor, started from KM 0 (latitude: 3.052355oN, 101.502278oE) to KM 9 (latitude: 3.074784oN, 101.555472oE). It is categorized as a municipal road, under the responsibility of Majlis Bandaraya Shah Alam (MBSA). The entire highway spans 9.3 km in length and the focused area are KM 0 - KM 9.

The site was chosen because it is listed as one of the black spot location in Selangor in 2015 by the Road Safety Department of Malaysia. Also, the site was chosen because the road signboard stating the road is the accident prone area is clearly displayed.

In this study, two approaches were practiced in collecting the accident data. Data was collected in the form of historical and from the site itself. Both sets of data are important to enable further analysis to be conducted.

After the analysis of all the traffic data, more sophisticated techniques were applied. The sophisticated techniques are the predictive analyses which are Multiple Linear Regression (MLR) and Artificial Neural Network (ANN) models. Each technique produced a model which predicts the accident frequency on current year based on the historical data from the past three years, 2013 to 2015 and also the traffic data which were classified as contributing factors and independent variables.

III. RESULTS AND DISCUSSION

This chapter is concerned to determine the final output for the prediction analysis, which uses two models, namely Multiple Linear Regression and the Artificial Neural Network thus comparing the effectiveness of the two predictive models from the errors obtained.

A. Accident Prediction Models

In the modern world, accident prediction models have been used for predictive analysis rather than conventional method. Therefore, it is important to acquire some knowledge on these models so that one can know which models are suitable according to particular situation. In this study, two models have been applied, which are the Multiple Linear Regression (MLR) and Artificial Neural Network (ANN). Both methods are applied and modelled to determine which model is most suitable for the accident predictive analysis.

Table 1. Comparison in R and MSE between MLR and ANN.

Model	Correlation coefficient (R)	Mean Square Error (MSE)
Multiple Linear Regression (MLR)	0.513	39,770.86
Artificial Neural Network (ANN)	0.935	867.309

From the Table 1 of comparison above, it is shown that the variables used in Artificial Neural Network are more correlated as compared to those in Multiple Linear Regression. Higher correlation indicates stronger relationship between independent variables. In this case, the relationship between hourly traffic volumes, 85th percentile speed, vehicles' gap and number of access point in Artificial Neural Network are stronger than relationship between hourly traffic volume and 85 th percentile speed in Multiple Linear Regression.

Higher correlation coefficient reduces the mean square error produced. This can be shown in the table above, where higher correlation coefficient in Artificial Neural Network results in lesser mean square error produced, thus showing greater accuracy in terms of prediction. Also, the accuracy of prediction can be analysed by logic, considering the value of historical number of accidents and the predicted values obtained from MLR and ANN, as shown in the Table 2 below.

Table 2. Comparison in R and MSE between MLR and ANN.

Section (km)	Actual			MLR (predicted)	ANN (predicted)
	2013	2014	2015	2016	2016
0	267	284	266	448	115
1	47	40	48	247	37
2	18	67	89	313	102
3	12	5	0	33	10
4	0	0	1	209	35
5	0	0	0	163	22
6	0	0	0	66	21
7	2	0	0	58	80
8	0	0	0	25	
9	37	39	10	86	
Total	383	435	414	1648	422

B. Accuracy of Prediction Model

In this section, the overall mean square error (MSE) and the correlation coefficient (R) between two models, Multiple Linear Regression (MLR) and Artificial Neural Network (ANN) were compared to determine respective accuracy, as shown in Table 1. The model which shows lesser MSE and higher R can be described as the most suitable model in this study. From the results shown above, it was clearly stated that MSE for ANN is much lesser than MLR and R for ANN is much more better which is near to value "1" compare to MLR.

IV. CONCLUSION

The prediction models used in this study have shown some advantages and disadvantages. Previous studies contradict the advantages between two models used, the Multiple Linear Regression (MLR) and Artificial Neural Network. For instance, two different researchers might have their respective stand on which model is the best. Therefore, it is concluded that predictive analysis based on Artificial Neural Network (ANN) is more suitable to be applied in this study rather than Multiple Linear Regression model. The conclusion is based on the criterion of the correlation coefficient and the mean square error produced by both models, in which the results from Artificial Neural Network (ANN) shows greater correlation and lesser mean square error, thus meeting all the criteria stated and ensure higher accuracy as compared to the other model, which is Multiple Linear Regression. The criteria stated supports the study by [15], whom stated that these two elements were chosen to determine the accuracy because they were adopted as criterion in for such analysis.

V. REFERENCES

[1] Road Transport Department Malaysia. Retrieved Mac 5, 2016, from http://www.jpj.gov.my/web/guest/pendaftaran-kenderaan-perdagangan.

[2] Mustakim, F., Aziz, A., and Samad, A. (2008). Blackspot Study and Accident Prediction Model Using Multiple Liner Regression. Transportation, 121-136. Retrieved from http://civil.neduet.edu.pk/ICCIDC-I/Conference Proceedings/Papers/013.pdf

[3] Oyedepo, O., and Makinde, O. (2010). Accident Prediction Models for Akure-Ondo Carriageway, Ondo State Southwest Nigeria; Using Multiple Linear Regressions. African Research Review, 4(2), 30-49. Retrieved from http://www.ajol.info/index.php/afrrev/article/view/58286

[4] Ramli, M. Z.,. (2011). Development of accident prediction model by using artificial neural network, Master thesis. Universiti Tun Hussein Onn Malaysia.

[5] Kamba, A. N., Rahmat, R. A. O. K., & Ismail, A. (2007). Why Do People Use Their Cars: A Case Study In Malaysia. Journal of Social Sciences, 3(3), 117-122. http://doi.org/10.3844/jssp.2007.117.122

[6] Mustakim, F. (2006). Treating Hazardous Location at Federal Route 50. Master of Science (Construction Management) Thesis. Universiti Teknologi Malaysia.

[7] Tian, R., Xiang, Q., & Hu, S. (2013). Research on Traffic Safety of Freeway Upgrade Section. Procedia - Social and Behavioral Sciences, 96(Cictp), 548-556. http://doi.org/10.1016/j.sbspro.2013.08.063

[8] Odumosu, A. O., (2005) Demeanour Transposition as Strategy for Traffic Accident Reduction in Nigeria: Case Study of Niger State, Nigeria, Proc. 13 th International Conference on Road Safety on Four Continents, Warsaw, Poland, p. 18

[9] Ogwueleka, F. N., Misra, S., & Ogwueleka, T. C. (2014). An Artificial Neural Network Model for Road Accident Prediction : A Case Study of a Developing Country. Acta Polytechnica Hungarica, 11(5), 177-197.

[10] Ali, G. A., & Tayfour, A. (2012). Characteristics and Prediction of Traffic Accident Casualties In Sudan Using Statistical Modeling and Artificial Neural Networks, 1(4), 305-317.

[11] Bastos, Y. G. L., de Andrade, S. M., Soares, D. A., & Matsuo, T. (2005). Seat belt and helmet use among victims of traffic accidents in a city of Southern Brazil, 1997- 2000. Public Health, 119(10), 930-932. http://doi.org/10.1016/j.puhe.2005.01.008

[12] Oyedepo, O., & Makinde, O. (2010). Accident Prediction Models for Akure-Ondo Carriageway, Ondo State Southwest Nigeria; Using Multiple Linear Regressions. African Research Review, 4(2), 30-49. Retrieved from http://www.ajol.info/index.php/afrrev/article/view/58286

[13] Abdelwahab, Hassan T., and Mohamed A. Abdel-Aty. "Development of artificial neural network models to predict driver injury severity in traffic accidents at signalized intersections." Transportation Research Record: Journal of the Transportation Research Board 1746, no. 1 (2001): 6-13.

[14] Akgüngör, A., & Doğan, E. (2009). An application of modified Smeed, adapted Andreassen and artificial neural network accident models to three metropolitan cities of Turkey. Scientific Research and Essay, 4(9), 906-913. Retrieved from http://www.researchgate.net/publication/228512198_An_application_of_modified_Smeed_adapted_Andreassen_and_artificial_neural_network_accident_models_to_three_metropolitan_cities_of_Turkey/file/3deec52b9692eb6454.pdf

[15] Hejmanowski, R., & Witkowski, W. T. (2015). Suitability assessment of artificial neural network to approximate surface subsidence due to rock mass drainage. Journal of Sustainable Mining, 14(2), 101 -107. http://doi.org/10.1016/j.jsm.2015.08.014