USE OF ARTIFICIAL NEURAL NETWORK ANALYSES ON INJURY PREVENTION SYSTEM

Phyu Win Nwe[1], Aye Thida Soe[2], Aye Aye Thinn[3]
Faculty of Computing, University of Information Technology
phyuwinnwe.pwn@gmail.com

ABSTRACT

The number of accidents on a given highway section during a certain period of time is probabilistic in nature and is a non-negative integer. Despite the fact that accidents are random and unpredictable at micro-level, statistical models can predict reliable estimates of expected accidents by relating aggregates of accidents to various explanatory measures. The research paper deals with injury prediction in traffic accident analysis which apply most commonly statistical predictive model for injury severity. This study aim to analyze categorical variable based on Artificial Neural Network (ANN). The data were acquired from the Government Digital Service (GDS), UK as categorical variables.

KEYWORDS

categorical variables, ANN, Injury severity

1. INTRODUCTION

Artificial Neural Network (ANN) systems have been applied in different information technology problems, such as traffic in communication and transportation engineering. ANN are employed for modeling the relationship that exist among driver injury severity and crash causes or factors that have to do with the driver, vehicle, roadway and the environment characteristics. Most traffic accident prediction models are based on statistical regression techniques. Smeed [1] studied the calculated number of fatally injured person in the accidents and compared the accident rates in different countries. Many communication traffic and transportation engineering problems have been solved using artificial neural network.

This study aim to explain the use of neural networks in the modeling of the number of persons fatally injured in motor vehicle accidents in dataset of the UK. In this paper, the development an Artificial Neural Network (ANN) model was carried out for the examination and prediction of accidents rate using UK as a case study. In the design of the system, accidents are selected and used as model parameters. We used the sigmoid and linear functions as activation function with the feed forward propagation algorithm. By analyzing the performance evaluation of the model, we found that the ANN model is better than other statistical methods in categorical analysis.

The paper is organized in the following section. In section 2 describes the related methodologies of proposed system and results and discussion are explained in section 3.

2. METHODOLOGIES

2.1 ARTIFICIAL NEURAL NETWORK (ANN)

The most common feed-forward neural network for classification has the following form:

$$p = \sigma(\sum_{i=1}^{l} \lambda_i \sigma(x.w_i)),$$

Where l is the number of hidden units, σ is the sigmoidal function given by $\sigma(x) = 1/(1 + \exp(-x))$, x are the inputs (covariates) and w are the weights (parameter) attached to each neuron. The design of the input includes an intercept term (often called "bias" in Neural Network lingo) so that $x.w_i = \sum_k x_k w_{ik} + w_{i0}$.

In term of log odds, the common feed forward network can be written as

$$\log\left(\frac{p}{(1-p)}\right) = \sum_{i=1}^{l} \lambda_i \sigma(x.w_i)$$

This is a nonlinear model for the effect of the inputs on the log odds, as each projection $x.w_i$has a nonlinear effect on the output mediated through the sigmoidal function.

In a manner similar to the interpretation of logistic regression, we study the effect of an infinitesimal change in variable $x_j$on the logit transform of the probability. Since p(x) is a smooth function of x, it is meaningful to examine the derivative

$$\frac{\vartheta}{\vartheta_{xj}} \log(\frac{p(x)}{1-p(x)} = \sum_{i=1}^{l} \lambda_i \sigma'(x.w_i) w_{ij}$$

In logistic regression, the effect of each covariate $x_j$on the log odds is given by the individual weights $w_j$since the odds are expressed as a linear combination of the inputs. The effect of each covariate $x_j$for the neural network model is given by what we term a generalized weight:

$$\widetilde{w}_j(x) = \sum_{i=1}^{l} \lambda_i \sigma'(x.w_i) w_{ij}$$

Thus, in neural network modeling, the generalized weights have the same interpretation as weights have in logistic regression, i.e., the contribution to the log odds.

2.2 Type of Variables

This section describes the characteristics of variables and explains how to use them in this study. First, it explains how variables can be characterized as either categorical or continuous, and second, it illustrates the role of response and explanatory variables. (Cohen, 1983) briefs descriptions of types of variables are presented:

Binary variable: Observations (i.e., response variables) that occur in one of two possible states, often labelled 0 or 1 (e.g., "improved or not improved" or "completed task or failed to complete task")

Categorical or Nominal Variable: Usually an explanatory or predictor variable that contains values indicating membership in one of several possible categories; categories are often assigned numerical values used as labels (e.g., 0, male; 1, female)

Continuous or Interval Variable: A variable not restricted to particular values (other than limited by the accuracy of the measuring instrument); equal size intervals on different parts of the scale assumed, if not demonstrated

Discrete Variable: Variable having only integer values (e.g., number of trials need by a student to learn a memorization task)

Ordinal Variable: A variable used to rank a sample of individuals with respect to some characteristics, but differences (i.e., intervals) and different points of the scale are not necessarily equivalent (e.g., anxiety might be rated on a scale "none, mild, moderate, and severe", with corresponding numerical values of 0, 1, 2, 3)

Outcome Variable: The presumed effect in a non-experimental study

Ratio Variable: Interval variables, but with the added condition that 0 of the measurement indicates that there is none of that variable (e.g., height, mass, distance and many more)

3. RESULTS AND DISCUSSION

Neural networks used in predictive application, such as the multilayer perception (MLP) and radial basis function (RBF) networks, are supervised in the sense that the model-predicted results can be compared against known values of the target variable. The term neural network (multilayer perception) is applied to the traffic accident data analysis of UK. The multilayer perceptron procedure produces a predictive model for one or more dependent variables based on the values of the predictor variables.

Table 1: Case Processing Summary

Case Processing Summary

		N	Percent
Sample	Training	34	66.7%
	Testing	17	33.3%
Valid		51	100.0%
Excluded		0	
Total		51	

Table 1 describes the case processing summary for how to classify the training and testing phase of traffic accident data analysis. Table 2 shows the factors information, number of hidden layer and output layer which are used in the categorical data analysis.

Table 2: Network Information

Network Information

Input Layer	Factors	1	Road Surface
		2	Lighting Conditions
		3	Weather Conditions
		4	Casualty Class
		5	Sex of Casualty
		6	Age of Casualty
		7	Type of Vehicle
		8	Number of Vehicles
	Number of Units[a]		22
Hidden Layer(s)	Number of Hidden Layers		1
	Number of Units in Hidden Layer 1[a]		1
	Activation Function		Hyperbolic tangent
Output Layer	Dependent Variables	1	Casualty Severity
	Number of Units		3
	Activation Function		Softmax
	Error Function		Cross-entropy

a. Excluding the bias unit

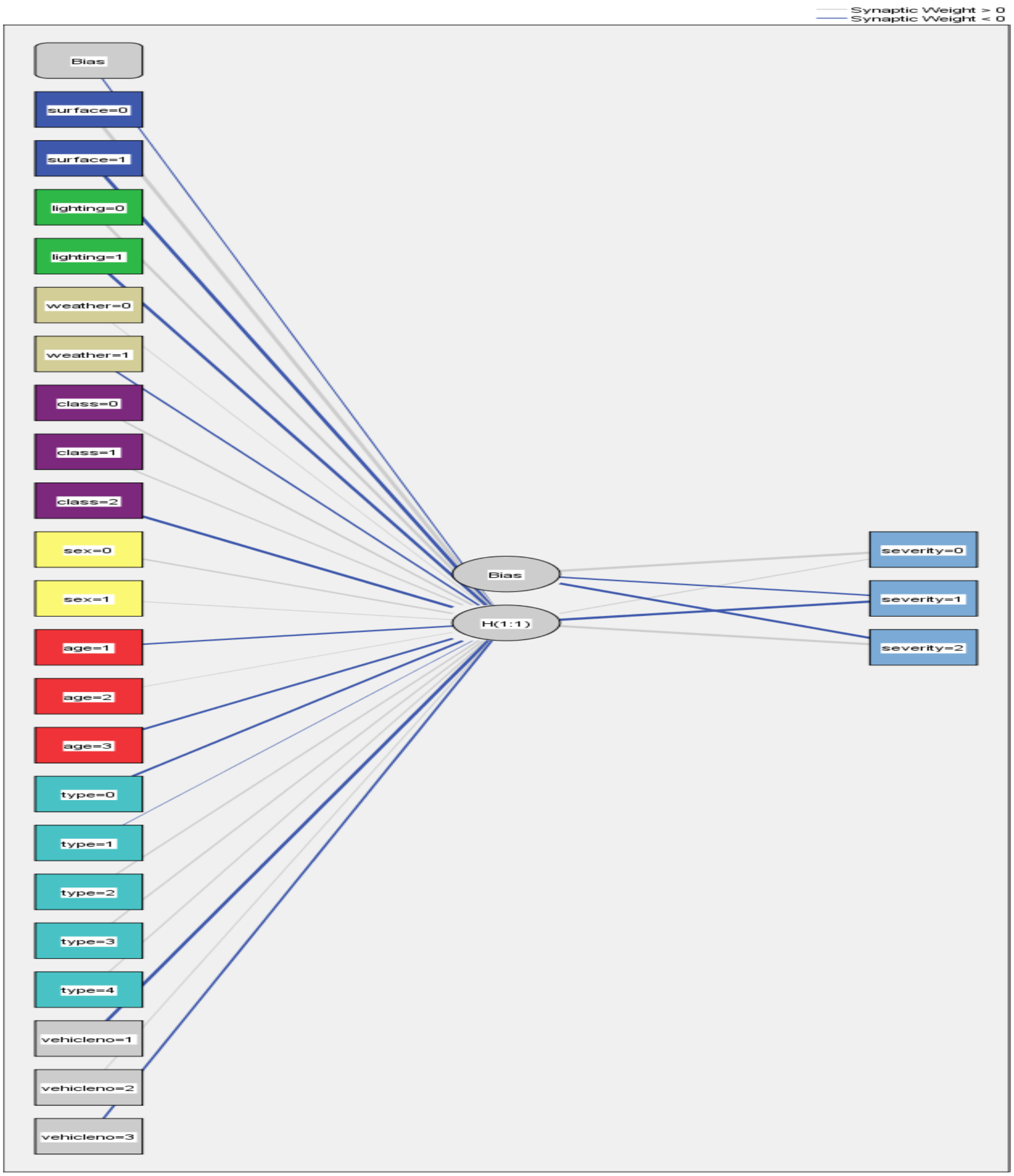

Figure 2: Neural Network of Traffic Accident Analysis

Figure 2 shows the structure of feed forward architecture because the connections in the network flow forward from the input layer to the output layer without any feedback loops. In this figure 2, input layer contains the predictors, hidden layer contain unobservable nodes, or units of accident data. That is a function (hyperbolic tangent) of predictors which depends in part upon the network type and in part upon user controllable specifications. The output layer contains the responses variable. This is categorical variable with three categories of injuries, (fatal, serious and slight).

As shown in the figure above, “Dry surface, Darkness: street lighting, fine without high winds, pedestrian, age group(1-40yr), Motorcycle over 50cc and up to 125cc” have more synaptic weight.

3.1 Measurement of Model Performance in Neural Network Structure (Architecture)

The following model summary table displays information about the results of the neural network training. Here cross entropy error is displayed because the out layer uses the softmax activation function. This is the error function that the network tries to minimize during training. Moreover, the percentage of incorrect prediction is equivalent to 17.025% in the training samples. Therefore, percentage of correct prediction is nearer to 77% that is acceptable. If any dependent variable has scale measurement level, then the average overall relative error (relative to the mean model) is displayed. On the other hand, if the defined dependent variables are categorical, then the average percentage of incorrect predictions is displayed.

Table 3: Model Summary

Model Summary

Training	Cross Entropy Error	17.025
	Percent Incorrect Predictions	23.5%
	Stopping Rule Used	1 consecutive step(s) with no decrease in error[a]
	Training Time	0:00:00.01
Testing	Cross Entropy Error	12.107
	Percent Incorrect Predictions	35.3%

Dependent Variable: Casualty Severity

a. Error computations are based on the testing sample.

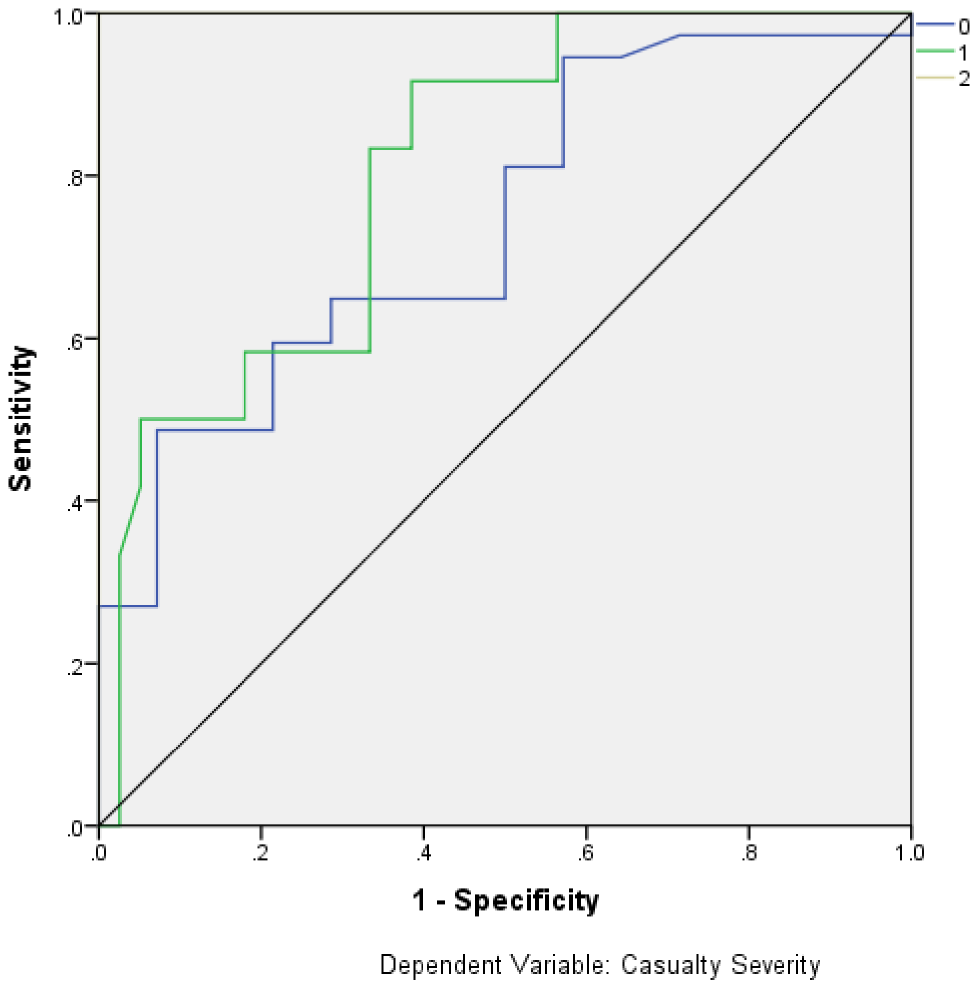

Figure 2: ROC curve for Traffic Accident Analysis

Figure 2 displays an ROC (Receiver Operating Characteristic) curve for each categorical dependent variable of injury severity. In this analysis, there are three categories of dependent variables; therefore each curve treats the category at issue as the positive state versus the aggregate of all other categories. This has the highest predictive power by using 7 independent variables for estimating the injury severity in UK traffic accident.

4. CONCLUSION

In this study, the factors which cause accidents have been investigated, for providing road safety, and accident prediction models which include relations between these factors have been established. In conclusion, it can be said that, in this model most of the impact from different factors can be seen on “fatality and serious” injury. These are the factors which will help prediction “fatality and serious” injury most efficiently and the rest will help in explaining “minor injury”.

On the contrary, it is important to note that the success of the predictive model largely depends on the explanatory variables selection to be used as inputs of the model.

REFERENCES

[1] Smeed, R. J., (1949) Some Statistical Aspects of Road Safety Research. Journal of Royal Statistical Society Series A 112, pp. 1-34

[2]Geman,S.,Bienenstock, E.,andDoursat,R.,Neuralnetworksandthebias-variancedilemma.Neural Computation, 4 (1992)1-58.

[3] Hansen,L. K. andSalamon,P., Neuralnetworkensembles.IEEETransactionson PatternAnalysisandMachineIntellignce, 12(10)(1990)993-1001.