© 2018 IJSRSET | Volume 4 | Issue 8 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
Themed Section : Engineering and Technology

An Artificial Neural Network Model For Road Accident Prediction

Rahul Purohit[1], Reshma Dayma[2]

[1]M.E(Computer engineering) Student, Computer Engineering Department, L.D College Of Engineering, Ahmedabad, Gujarat, India

[2]Assistant Prof., Computer Engineering Department, L.D College Of Engineering, Ahmedabad, Gujarat, India

ABSTRACT

Road related accidents are considered one of the most serious problems in the modern world as traffic accidents cause serious threat to human life worldwide. Prediction of future traffic accidents is therefore of utmost importance in order to appreciate the magnitude of the problem and speed up the decision making towards its alleviation. This research presents an accident prediction model using artificial neural network (ANN) approaches. This neural models use the traffic accident data with the OBD(on board diagnostic) as training samples and testing the models

Keywords: OBD Device, Road Accident prediction, Artificial Neural Network.

I. INTRODUCTION

An Artificial Neural Network (ANN) is an information processing paradigm that's inspired by the way biological nervous systems, comparable to the brain, process information. The key part of this paradigm is that the novel structure of the knowledge process system. it's composed of an outsized range of extremely interconnected processing elements (neurones) operating in unison to solve specific issues. ANNs, like folks, learn by example. an ANN is designed for a selected application, such as pattern recognition or knowledge classification, through a learning method. Learning in biological systems involves adjustments to the colligation connections that exist between the neurons

Neural networks take a special approach to problem solving than that of standard computers. standard computers use an algorithmic approach i.e. the computer follows a group of instructions in order to solve a problem. Unless the particular steps that the computer needs to follow are known the computer cannot solve the problem. That restricts the problem solving capability of conventional computers to problems that we tend to already understand and skills to resolve. however computers would be so much more useful if they might do things that we do not precisely skills to try and do.

Neural networks method info in an exceedingly similar manner the human brain will. The network consists of an outsized range of extremely interconnected process elements(neurons) operating in parallel to unravel a selected drawback. Neural networks learn by example. they can not be programmed to perform a selected task. The examples should be chosen rigorously otherwise helpful time is wasted or perhaps worse the network could be functioning incorrectly. The disadvantage is that as a result of the network finds out a way to solve the matter by itself, its operation are often unpredictable.
On the opposite hand, typical computers use a cognitive approach to problem solving; the approach

the matter is to resolved should be best-known and expressed in tiny unambiguous directions. These directions are then converted to a high level language program and so into machine code that the computer will understand. These machines are entirely predictable; if something goes wrong is due to a software or hardware fault.

On-board medicine (OBD) is an automotive term referring to a vehicle's self-diagnostic and reportage capability. OBD systems offer the vehicle owner or repair technician access to the status of the various vehicle subsystems. the number of diagnostic info obtainable via OBD has varied wide since its introduction within the early Eighties versions of on-board vehicle computers. Early versions of OBD would merely illuminate a malfunction indicator light-weight or "idiot light" if a problem was detected however would not give any info on the nature of the matter. trendy OBD implementations use a uniform digital communications port to supply period of time knowledge additionally to a consistent series of diagnostic bother codes, or DTCs, which permit one to chop-chop determine and remedy malfunctions among the vehicle.

II. DIFFERENT ANN MODEL APPROACHES USED FOR THE ROAD ACCIDENT PREDICATION.

Prediction of Road Traffic Accidents in Jordan[4] concerned a sequence of steps as follows:
Data collection section including the subsequent input data:

- ✓ V: range of registered vehicles
- ✓ P: population
- ✓ L: total length of paved roads
- ✓ G: the gross domestic product

The data is split into 3 sets; training data.(about seventieth of the entire data), validation data (about V-day of the total data), and testing data (about15% of the full data). Training, validation and testing of the network was performed using MATLAB.

The training method includes the following operations:

- ✓ Setting initial values for weights
- ✓ Evaluating the output supported initial weights.
- ✓ Measuring the error (mean square error or any function to calculate the error)
- ✓ Adjusting the weights using rate of learn(usually small value similar to 0.01)
- ✓ The weights still be changed as every error is computed. If the network is capable and also the learning rate is about properly, the error is eventually driven to zero.

In the validation phase, no adjustment happens to the weights. Validation is important to measure the performance of the network model wherever the predicted values are compared with the actual as given by the validation knowledge. This method may be integrated with training method to boost the performance of the model Through the testing method, the predicted values are compared with the input values using testing data that wasn't utilized in training or validation method. Again, no adjustment happens to the weights as a results of the ANN model they predicted {the range| the amount| the quantity} of accident per annum based on the given number of registered vehicles, population, total length of paved roads and also the gross domestic product.

In case study of road accident prediction of erzurum, turkey [5] the collected data of the quantity of accidents covered a period of eight years from 2005 to 2012 and relate to the road network of the Province of Erzurum in eastern Turkey. The traffic accident reports data utilized in this study are 7,780 complete accident reports that were collected from the Directory of Erzurum Traffic Region. every accident report has numerous information such as the date, accident location, pavement type, vehicle type, driver's gender, driver's age, road surface condition, the day and time, weather condition, day or nighttime, the number of deaths, the number of injured persons, the number of involved vehicles, and the number of damaged vehicles. In addition to these data, geometric characteristics of the highway such as AADT, the degree of horizontal and vertical curvatures in each section, lane, median, and shoulder widths were collected from 12th Highway Regional Directorate of Erzurum. Here, there are eight input variables containing 31 neurons that are the input variables representing the potential risk factors for accidents.

Before the training of the network, both input and output variables were normalized within the range 0-1 using a minimax algorithm. Categorical columns were automatically encoded during the data preprocessing using One-of-N method by Alyuda Neuro Intelligence program. The One-of-N encoding means that a column with N distinct categories (values) is encoded into a set of N numeric columns, with one column for each category. For example, for the Capacity column with values “Low”, “Medium” and “High”, “Low” will be represented as {1,0,0}, Medium as {0,1,0}, and High as {0,0,1}. The minimum and maximum of the dataset were found and scaling factors selected so that these were mapped to the desired minimum and maximum values.

III. PROPOSED ALGORITHM

The choice of an appropriate input vector that will allow an ANN to clearly map to the desired output vector is not a trivial task]. One of the most important tasks of the modeler is to find out which parameters are influencing the system under work. A firm knowledge of this OBD2 device system is therefore necessary, because this will allow the modeler to make better selection regarding the input variables for proper mapping. This will, on the one hand, help in avoiding loss of data(e.g. if key input variables are omitted), and, on the other hand, prevent unnecessary inputs from being used to the model, which can result in decreasing network performance. many applications have proven the usefulness of a trial-and-error procedure in determining whether an ANN can extract information from a variable.

Since empirical modelling is data-driven modelling the necessary of data quality is not to be neglected. The first subsection discusses data quality, quantity and representatively. Data requirements The input-output patterns, which are used to make the network learn during the training phase, are to be chosen in such a way that a good ANN model will be able to abstract enough information from them to manage in the networks operational phase.

The resolution of data has to be in proportion to the system under investigation. Within the context of lumped models (where spatial variability is often ignored) the time resolution is often the only consideration.

SensApp is a platform to support sensor based application. It is developed by SINTEF (IKT division, NSS Department, MOD research group).As a basis, SensApp provides four essential services to support the definition of IoT applications. The Registry stores metadata about the sensors (e.g., description and creation date). The Database service stores raw data from the sensors using a MongoDB database. The Notifier component sends notifications to third-party applications when relevant data are pushed (e.g., when new data collected by air quality sensors become available).This gives the acces to the jason file of the car's data with parameters like "longitude, latitude, GPS speed, Engine RPM, OBD_Speed, Liter per 100km instance, Engine co2 level, engine coolant, GPS altitude".

The following steps are taken when the data are in proper format.

Step 1 : Data collection phase

Step 2 : Data divided in 70% for training, 30% for testing.

Step 3 : Input Hidden Layer use the ReLU transfer function

Step 4 :Output Hidden Layer SoftMax or Linear transfer function.
Step 5 : Setting initial values for weights
Step 6 : Evaluating the output based on initial weights
Step 7 : Measuring the error. (mean square error or other.)
Step 8 : Adjusting the weights using rate of learn
Step 9 : The weights continue to be modified as each error is computed. If the network is capable and the learning rate is set correctly, the error is eventually driven to zero.

IV. ANALYSIS OF PROPOSED METHOD

Implementation of proposed algorithm is carried out using Anaconda python. After the converting the data of json format to csv format.

The different prediction models for the road accident related issue and their comparison are as follow.

Table 1

	Data set	ANN Model used	Prediction about
Model 1	number of registered vehicles, population, total length of paved roads, the gross domestic product	MLP	Accident rate in given criteria in
Model 2	years, highway sections, section length (km), annual average daily traffic (AADT), the degree of horizontal curvature, the degree of vertical curvature, traffic accidents with heavy vehicles (percentage), and traffic accidents occurred in summer (percentage)	kNN	Accident rate in given criteria
Model 3	per capita GDP, numbers of vehicles and highway mileage	GA-BP learning algorithm	Accident rate at given criteria
Our model	Longitude, latitude, GPS altitude, GPS speed.	MLP algorithm	Accident probability at given criteria and location

As no other model used to predict the road accident probability of the car during given car's situation. rather most of the ANN model build around the prediction about the percentage of accident rate on given criteria, we can't compare our result with this models.

So we applied the different ANN models to our data set and gathered the result.
Result comparison with different classifier:

Table 2

Classifier	Best result achieved from 100% accuracy
1) GaussianProcessClassifier	82.23
2) KNeighborsClassifier	89.53
3) DecisionTreeClassifier	84.12
4) RandomForestClassifier	89.23
5) MLPClassifier	92.25

As we can see the MLP classifier has the best result with the 92.25 accuracy.so the prediction about the accident prone condition of the car of given GPS altitude, GPS speed and the latitude and longitude gives the predictable result.

V. CONCLUSION

All the previous ANN model for the prediction of road accident used the data from the traffic sensor and the historical accident scenario based data. There is no ANN model which are predicating accident using the vehicles data. we can use the feature rich data from the OBD device of vehicles which can give the exact information of the vehicle at time of the accident and before of It, and then can apply the different NN model to predicate the road accident more accurately.

Here we are only working on the data which are collected during the time span of the year, if we apply this prediction model on the real time streaming data from the all the transportation medium with the OBD device, we can have more accurate and practically applicable prediction result that can be used to driver's attention before accident. In future scope of this work, we can add more data frames regarding the environmental condition, traffic data, road sensor data to predicate more relevant to the exact time and surrounding condition.

VI. REFERENCES

[1]. Nikhil Buduma, 2015. Fundamentals of Deep Learning.2th Edn., Sebastopol, CA., ISBN: 978-1-491-92561-4

[2]. Prediction of Road Traffic Accidents in Jordan using Artificial Neural Network (ANN) Khair S. Jadaan, Muaath Al-Fayyad, and Hala F. Gammoh University of Jordan /Civil Engineering Department, Amman , Jordan. doi: 10.12720/jtle.2.2.92-94

[3]. An artificial neural network model for highway accident prediction: a case study of erzurum, turkey. Promet – Traffic&Transportation, Vol. 27, 2015, No. 3, 217-225

[4]. Macro prediction model offroad traffic accident based on neural network and genetic algorithm QIN Liyan SHAG Chunfu DOl :10. 1109/ICICTA. 2009. 93

[5]. Aroui T, Koubaa Y, Toumi A (2008) „Magnetic Coupled Circuit Modelling of Inductionmachines oriented to Diagnostics", Leonardo Journal of sciences, Vol. 7, issue 13, pp.103- 121, ISSN: 1583-0233.

[6]. G.A. Gonçalves, T.L. Farias, "On-road measurements of emissions and fuel consumption of gasoline fuelled light duty vehicles", Eighth International Conference on Technologies and Combustion for a Clean Environment, Lisbon, Portugal, 2005.

[7]. Z. Wenhui, "Safety evaluation of traffic accident scene based on artificial neural network," ICICTA, China, vol. 1, Oct. 2009, pp. 408-410.

[8]. Y. C. Chiou, " An artificial neural network- based expert system for the appraisal of two-car crash accidents," Department of Traffic and Transportation Engineering and Management, Feng Chia University, Taiwan, July 2006.