RESEARCH ARTICLE OPEN ACCESS

Pedestrian Path Prediction with Neural Network

Gandrapu Gideon [1], Prof. K. Venkata Rao [2]
Department of Computer Science and Systems Engineering
Andhra University College of Engineering (A)
Visakhapatnam
AP - India

ABSTRACT
The deaths from traffic accidents as well as injuries have become very common everywhere, and this may be due to poor driving or bad weather conditions or may be due to pedestrian's negligence. In order to prevent these accidents and crashes pedestrian detection system is deployed in some cars and this system alone cannot be a solution to prevent accidents as this could be sometimes too late and accident could happen in the blink of an eye. So, in this paper, we address the problem of predicting the pedestrian location UK Car accidents dataset. Here we present a neural network algorithm which helps in making predictions on pedestrian location using other parameters which are some influencing factors for accidents. Firstly we train the algorithm with some data and then testing is made on the other part of the data. Finally we observe the M.S.E for this model.

Keywords :- Prediction, Neural Network, WHO, Road Accidents

I. INTRODUCTION
There has been increase in the road accidents everywhere this year which also resulted in the increase of death rate. Some reports from World Health Organization reveals Approximately 1.3 million people die each year on the world's roads, and between 20 and 50 million sustain non-fatal injuries. According to some report on road accidents analysis, the road accidents are the main issues especially for middle and lower middle level countries [1]. And from the last few years there is a major progression made in vehicle designing and also providing safety measures for pedestrians, this gave rise to a thought of providing one more prevention measure for pedestrian crashes. Many new technologies are in use today in order to prevent accidents but that doesn't always solve the problem of accidents and crashes as the pedestrian activities are always dynamic and this kind of problem needs a regular predicting model which can help in knowing the location of pedestrians in the streets and help the drivers. This also doesn't disturb the traffic and doesn't need any extra equipment fixed on the roads. In this paper, we also included some of the other influencing parameters which together provides a summary and overview of all the factors that how and among in which range these factors affect road traffic accidents. All these factors provides the road safety community and the traffic police with a better knowledge of road accidents and also helps in developing suitable methods which improves the road safety. Not only the common researched factors like speed, traffic flow, driver's age etc but also other factors related to environment like weather conditions and road surface also are the affecting factors for road accidents which are not always recorded or noted. World Health Organization says the newly adopted 2030 Agenda for Sustainable Development has set an ambitious target of halving the global number of deaths and injuries from road traffic crashes by 2020. Some analysis says if no prevention measures are taken for the road accidents by 2030 it will be the seventh main cause for human life loss [2]. This paper is implemented using regression technique, basically a regression is a parametric technique used to predict continuous (dependent) variable given a set of independent variables. It is parametric in nature because it makes certain assumptions (discussed next) based on the data set [12].

II. LITERATURE REVIEW
Joon-Young Kwak et al. [3] Illustrates the prevention method for pedestrian and vehicle accidents that happen during night time, which is a major problem everywhere. Many new driver assistant technologies are developed. These systems contains the cameras out of which far infrared cameras are given more importance as they have a property to capture the images and videos even in the lowest possible lights. Here they have used two different methods. The first one is to detect the pedestrian moments based on the season and weather conditions. They can be studied with the help of Weber-Frechners law. And the learning process is done using the Random Ferns.

Nicolas Schneider et al. [4] constructed a framework on Bayesian filters for pedestrian path detection. These are applied to different pedestrian movements. These pedestrian movements are measured using an external state-of-the-art stereo vision-based pedestrian detector. They also evaluated the accuracy of position estimation and path predicted. They also studied the importance of IMMs vs. the simpler single dynamical models.

R. Quintero et al. [5] developed pedestrian action based detector. They have used GPDM (Gaussian Process Dynamical Models) in their experiment. Here all the subjects are separately trained. Every action is individually classified here.

Wang et al. [6] researched on the different factors leading to accidents. Different parameters are discussed here so that the researchers whoever is planning to build a prevention method for road accidents can first study these factors and based on that they can introduce their methodology.

Sarah Bonnin et al [13] "worked on the behavior study of the pedestrian which crossing the roads. These system is about to give an alert signal to the driver whenever a pedestrian crosses the road. An early prediction is made in order to prevent the crashes on roads.

III. METHODOLOGY

A. *Problem Statement*

Here we are addressing the problem of predicting pedestrian path using neural network. For this study we have used Neural Network regression method of predictions of future data. Regression technique is a statistical technique which shows the relationship between the dependent variable and the independent variables. After fitting the model and here we also made a linear model parallel with the Neural Network model. M.S.E is observed for both after training and testing is done.

B. *Proposed System Architecture*

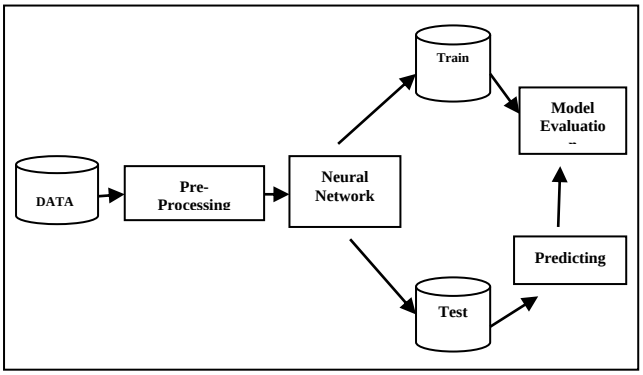

Fig. 1 System Architecture

Firstly the input which is the accidents dataset selected is given as input to the system. Then after in the preprocessing step missing values are checked. Here no missing values are found so no imputation is done. The dataset is then Scaled properly with min-max scaling in the intervals [0,1] or [-1,1] to produce a better results. Then splitting is done, 75% is used in training set and 25% is used for testing set. With the help of training set we trained the model at what parameters rise and fall will lead to what conditions and based on that we tested the model with the help of testing set.

IV. DATASET USED

A dataset consists of organized recorded data in blocked structure or record format that will be used by IBM mainframe operating systems. These datasets are uniquely named according their individual properties [14]. The dataset which is used here in this model is the UK Car Accidents 2005-2015 data which was collected by UK department of Transport, which was downloaded from kaggle [7]. This dataset actually contains three different datasheets with different categorized data individually, we have only included few columns and rows which are sufficient for our research work. This contains 11 columns and 785 rows. This has "dow" which is Day of the week, "noc" which is Number of casualities, "rt" which is road type, "jd" which is junction details, "nov" which is number of vehicles from accidents datasheet, "sl" which is the speed limit, "gen" which is the gender of casualty,"pl" which is Pedestrian Location, "sev" which is casualty severity from casualties datasheet. The last two parameters for this datasheet that are "ovrt" and "aov". These two parameters "ovrt" and "aov" has the information like the skid and overturn details and age classification recorded about the vehicles in the vehicle datasheet. After selecting these parameters we selected "pl" as our target variable which is needed to be predicted. We then divided our dataset 75% into training set and remaining 25% into testing set. Training dataset is used to train the algorithm for this particular data and testing dataset is to used to make predictions with the help of prior data, which is nothing but the training datset [15].

V. NEURAL NETWORK

Neural Network is the model built based on the human neural system. The flow of the Neural Network changes due to the inputs and outputs as it learns and processes the information [8]. The advantage of the usage of neural networks for prediction is that they are able to learn from examples and they catch the non-linear depencies and the disadvantage of Neural Network is they can be trained for certain period only [10]. Fitting the Neural Network is not easy or cannot be implemented without preprocessing work. Firstly we check for the missing values in the pre- processing step.

```
> set.seed(500)
> apply(data,2,function(x) sum(is.na(x)))
  dow   noc   nov    sl   gen   sev    rt    jd   aov
    0     0     0     0     0     0     0     0     0
overt    pl
    0     0
>
```

Fig. 2 Missing Values

From the fig 2 console, we can clear see there are no missing values in the data. So secondly in pre-processing we have done scaling process. There are many scaling procedure we can apply but here we have used min-max scaling with [0,1] or [-1,1] intervals. Then comes splitting dataset into two different sets one for training and another for testing. Here we have divide 75% in training and remaining 25% in testing. First we developed a linear model and we tested our data on the test set. Usually basic Neural Network has single perceptron model, but here we have taken 2 hidden layers with this configuration 10:5:3:1 where 10 is the inputs given, 5 and 3 are the neurons in the first and second hidden layers

and 1 is the output layer. First layer we have all the inputs given to the algorithm which are the “dow”, “noc”, “noc”, “sl”, “gen”, “rt”, “jd”, “aov”, “overt”. These are the predictor variables which are used to make predictions for our target variable which is “pl”.

The black lines in the fig 3 are the connections formed between the neurons in all the layers, blue lines which are also connected to neurons are the bias terms. And the numbers on these lines are the weights assigned to it. These weights are very important as a set of weighted inputs allows each artificial neuron or node in the system to produce related outputs [9].

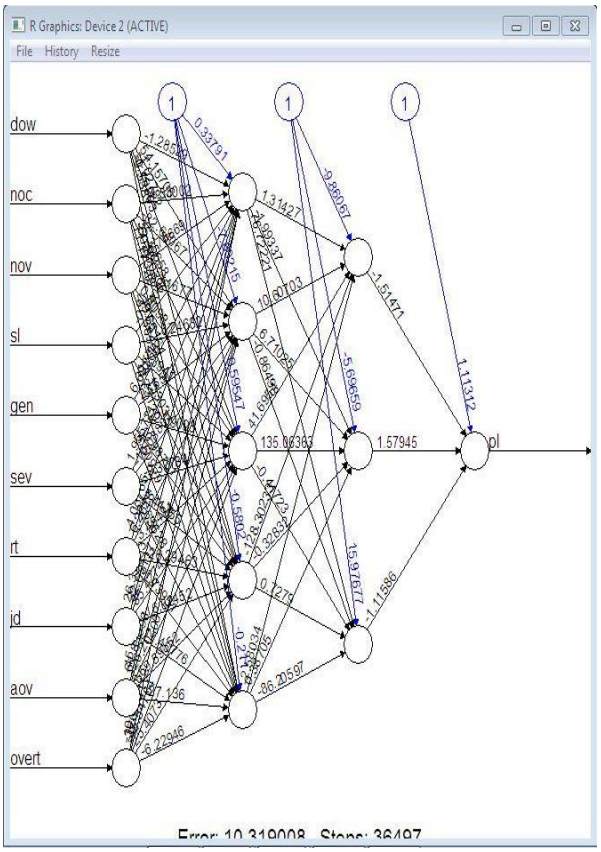

Fig. 3 Neural Network

After fitting the Neural Network model predictions are made on our target variable which is “pl” using the predictor variables. We then made predictions on the test data and for both the linear model and for the Neural Network model.

```
> MSE.nn <- sum((test.r - pr.nn_)^2)/nrow(test_)
> print(paste(MSE.lm,MSE.nn))
[1] "5.84019305520694 9.31481028911913"
>
```

Fig 4. M.S.E

From the above figure we can see the two mean square errors that is for both the linear model and for the Neural Net model.

Comparing both the mean square errors linear model has the lowest error rate for this data. This also depends on the splits made for training and testing data and also on other factors. The prediction made on any data using Neural Network can be trained for certain period of time and that learned information can be applied on the future data in that period of time [11].

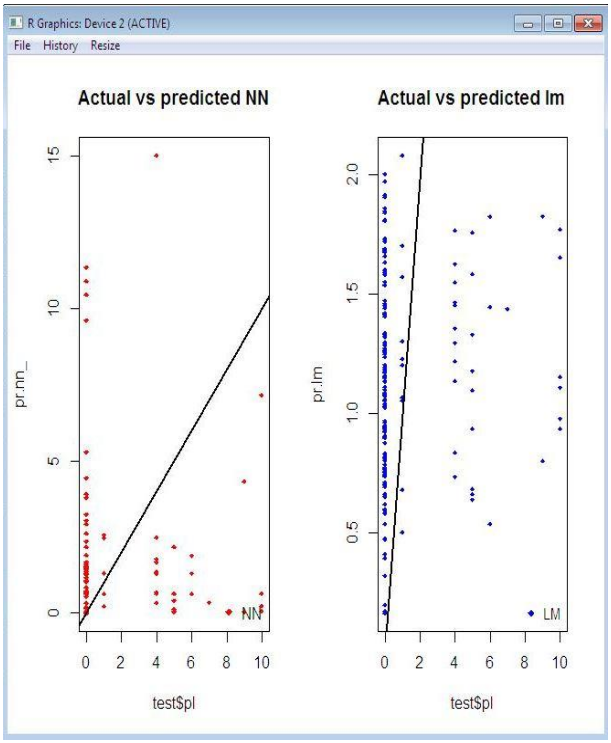

Fig. 5 Actual VS Predicted for Neural Network and Linear Model

Fig 5 clearly shows the actual values which were in the trained data and predicted values which were predicted using the test data for both normal linear model and for our Neural Network model. We can say it predicted very well as many points are found around the regression line. Fig. 6 Shows more clear view for both Linear model and Neural Network model in one graphical view. Red dots represent the Neural Network model predicted values and blue dots are the ones which are predicted using Linear Model.

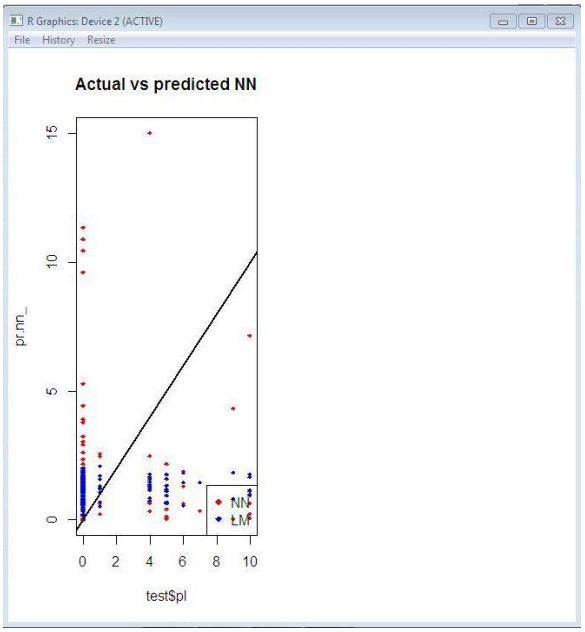

Fig. 6 Actual VS Predicted for both Neural and Linear model

VI. CROSS VALIDATION

Cross Validation is basically a very important step for predicting models. This step is nothing but the complete process is repeated for number of times and then calculating the M.S.E on average to know how far our model works well for this data.

Fig. 7 Average M.S.E

Here the average mean square error can be observed and there after every individual mean square error can also be seen. The bar represents that the network has completely compiled all the 10 times for the same data. And the bloxplot for this error can be observed from fig 8.

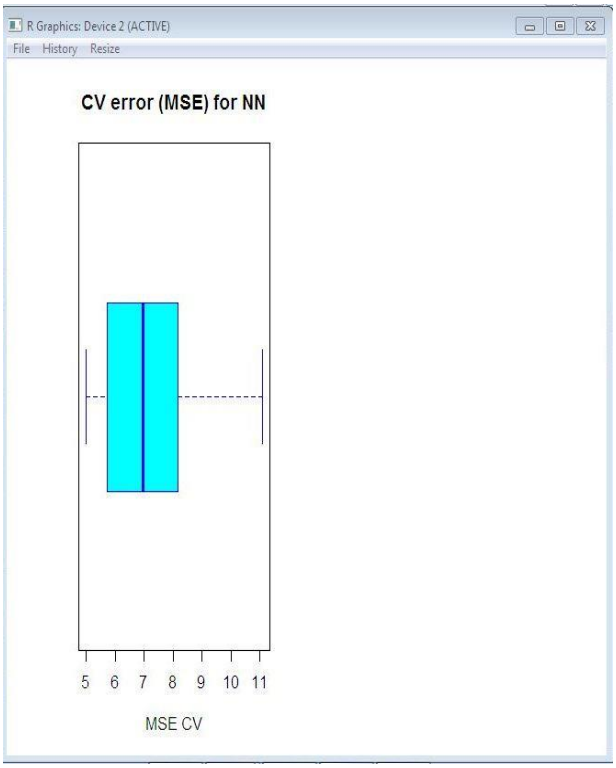

Fig. 8 Boxplot

VII. CONCLUSION

This paper proposes a pedestrian path prediction method using Neural Networks which first learns and then makes predictions on the dataset. Neural Network model has given a good performance. Manual guidance or arranging the cameras not always give a perfect solution to prevent accidents but sometimes a prediction which is made as in this paper may also help in preventing accidents. And this kind of predictions doesn't need any economical support or any additional equipments so these predictions are always helpful in preventing accidents. There are certain other parameters available to make more research work and this work definitely has a very good scope in future.

REFERENCES

[1] "Violence and Injury Prevention", [Online]. Available: https://www.who.int/violence_injury_prevention/road_safety_status/report/en/

[2] "Road traffic injuries", [Online]. Available: http://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries

[3] Joon-Young Kwak,Byoung Chul Ko,Jae Yeal Nam "Pedestrian Tracking Using Online Boosted Random Ferns Learning in Far-Infrared Imagery for Safe Driving at Night", Volume 18 Issue 1, January 2017 Pages 69-

81 [Online]. Available: https://dl.acm.org/citation.cfm?id=3084644

[4] Nicolas Schneider,Dariu M. Gavrila "Pedestrian Path Prediction with Recursive Bayesian Filters: A Comparative Study", LNCS, volume 8142, [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-642-40602-7_18

[5] R. Quintero, I. Parra, D. F. Llorca and M. A. Sotelo, "Pedestrian Path Prediction based on Body Language and Action Classification", [Online]. Available: http://www.isislab.es/raul/papers/ITSC2014.pdf

[6] Wang, Chao,Quddus, Mohammed A,Ison, Stephen, "Factors Affecting Road Safety: A Review and Future Research Direction", [Online]. Available: https://trid.trb.org/view/1129330

[7] "UK Car Accidents 2005-2015" [Online]. Available: https://www.kaggle.com/silicon99/dft-accident-data#Accidents0515.csv

[8] "Artificial Neural Network (ANN)", [Online]. Available: https://www.techopedia.com/definition/5967/artificial-neural-network-ann

[9] "Weight", [Online]. Available: https://www.techopedia.com/definition/33274/weight-neural-networks

[10] "Predictions Using Neural Network", [Online]. Avalaible: http://www.obitko.com/tutorials/neural-network-prediction/prediction-using-neural-networks.html

[11] "Neural Network Training", [Online]. Available: http://www.obitko.com/tutorials/neural-network-prediction/neural-network-training.html

[12] "Beginners guide to Regression Analysis and Plot Interpretations", [Online]. Available: https://www.hackerearth.com/practice/machine-learning/machine-learning-algorithms/beginners-guide-regression-analysis-plot-interpretations/tutorial/

[13] Sarah Bonnin,Thomas, H. Weisswange,Franz Kummert, Jens Schmuedderich, "Pedestrian crossing prediction using multiple context-based models", [Online]. Available: https://ieeexplore.ieee.org/document/6957720/metrics#metrics

[14] "Data Set", [Online]. Available: https://www.techopedia.com/definition/3348/data-set-ibm-mainframe

[15] "What is the difference between training data and testing data?", [Online]. Available: https://www.quora.com/What-is-the-difference-between-training-data-and-testing-data