Neural Computing and Applications (2020) 32:12417–12426
https://doi.org/10.1007/s00521-019-04695-8

ORIGINAL ARTICLE

Deep neural network-based predictive modeling of road accidents

Gyanendra Singh[1] · Mahesh Pal[2] · Yogender Yadav[2] · Tushar Singla[2]

Received: 30 August 2019 / Accepted: 18 December 2019 / Published online: 9 January 2020

Abstract

This work proposes to use deep neural networks (DNN) model for prediction of road accidents. DNN consists of two or more hidden layers with large number of nodes. Accident data of non-urban sections of eight highways were collected from official records, and dataset consists of a total of 2680 accidents. The data of 16 explanatory variables related to road geometry, traffic and road environment were collected from official records as well as through field studies. Out of a total of 222 data points of accident frequency, 148 were used for training and remaining 74 to test the models. To compare the performance of DNN-based modeling approach, gene expression programming (GEP) and random effect negative binomial (RENB) models were used. A correlation coefficient value of 0.945 (root mean square error = 5.908) was achieved by DNN in comparison with 0.914 (RMSE = 7.474) by GEP, and 0.891 (RMSE = 8.862) by RENB with the test dataset, indicating an improved performance by DNN in prediction of road accidents. In comparison with DNN, though lower value of correlation coefficient was achieved by GEP model, it quantified the effects of various variables on accident frequency and provided a ranked list of variables based upon their importance.

Keywords Gene expression programming · Deep neural network · Back-propagation neural network · Random effect negative binomial model · Accident prediction model · Road safety

1 Introduction

With a road network of 5.603 million km [1], India had a total of 4,64,910 road accidents in the calendar year 2017 leading to 1,47,913 causalities and 4,70,975 injuries [2]. Analysis of accident severity in terms of number of fatalities per 100 accidents suggests an increasing trend from 18.1% in 2003 to 31.8% in 2017 [3]. National and state highways comprise 4.94% of total road length only but account for 52.4% accidents, 63% fatalities and 55.3% injuries. Studies suggest that up-gradation of highways from two to four lanes has not resulted in safer travel [4].

As the real-time detailed driving data of road accidents are not available, accident prediction models are used as a useful tool to understand the causal factors of road accidents by engineers and planners and to suggest measures for road safety improvement. Accident prediction models try to relate the factors associated with accident occurrence by developing statistical relations that correlate various risk factors with the number of accidents occurring on a road section over a period of time. The occurrence of road accidents being a complex phenomenon, the possibility of making causal inferences based on accident prediction models depends strongly on how well the assumptions reflect the reality, what functional relationship was chosen and which method was adopted to overcome disturbing factors [5, 6]. The data and methodological issues involved in predictive modeling have been thoroughly discussed in the literature [5–8]. Indian scenario has added complexities due to heterogeneity of traffic [9], under-reporting of accidents [10, 11] and poor quality of available accident

✉ Gyanendra Singh
gyanendersingh.civil@dcrustm.org; singhgyan27@yahoo.in

Mahesh Pal
mahesh1965@gmail.com

Yogender Yadav
yogender105@gmail.com

Tushar Singla
tusharsingla06@gmail.com

1 Civil Engineering Department, DeenbandhuChhotu Ram University of Science and Technology, Murthal, Sonepat, Haryana, India

2 Department of Civil Engineering, National Institute of Technology, Kurukshetra, Haryana, India

data [12, 13]. Therefore, separate India-specific models are required to understand and quantify effect of various factors on accidents [14].

Parametric and nonparametric modeling techniques are two broad categories to predict road accidents. Parametric models (such as Poisson and negative binomial model) are statistical models which require prior assumptions about distribution of accidents and the relationship between accidents and factors affecting accidents [13, 15–21]. Thus, they are able to quantify effect of various influencing factors on accidents but their predictive power is seriously compromised if the assumptions are not correct. On the other hand, nonparametric models such as neural network [22], fuzzy logic [23], support vector machine models [24] and decision-tree based algorithms [25, 26] are also employed to predict accidents. The literature suggests that these algorithms when trained with a set of available data are able to predict accidents with reasonable accuracy but are not able to quantify the effect of influencing variables over accidents, and their predictive power is limited by data specifications.

Keeping issues related to road safety in mind, this work examines the effectiveness of deep neural network (DNN) [27] and gene expression programming (GEP) [28] model for prediction of road accidents and identification of the key risk factors. Within last few years, several studies reported the use of DNN in various civil engineering problems and found them performing well in comparison with the existing modeling approaches [29–37]. DNN can better represent highly varying and complex high-dimensional functions in comparison with conventional neural networks [38]. Detailed literature reviews suggest no application of DNN in accident prediction so far. GEP has also been used successfully in various fields [39–42]. The advantages of GEP over other nonparametric algorithms include providing a list of best explanatory features used to create the model as well as a relation between the output and explanatory variables. An exhaustive literature review suggests that the use of GEP has also not been explored for road safety and predictive modeling of road accidents.

To compare the results of DNN and GEP models, random effect negative binomial (RENB) regression model was used. RENB model is an advanced version of negative binomial model which takes into account unobserved heterogeneity by considering temporal and spatial correlation most often present in the accident data which is collected on interconnected roads repeatedly for a number of years [17, 43].

2 Materials and methods

2.1 Dataset

The data used for this study were collected over a number of sections of eight highways in Haryana (India). The data for road geometry and road environment were collected through field visits along various highways. Traffic volume data for different roads were collected from Toll plazas, various detailed project reports (DPRs) prepared by National Highway Authority of India (NHAI) and from traffic registers of Haryana public works department and supplemented by 16–24 h traffic count as per the requirement. The data of spot speeds were collected from DPRs and missing data by using radar gun. On the basis of the collected data, road sections having similar traffic, road width, median width and shoulder width were identified. Total road length of various highways covered in this study was more than 308 km and was divided in 68 uniform sections of varying length as shown in Table 1. Road accidents data for a period of 2–9 year on different sections (based upon availability) were collected from police records (Table 1). The accident data include all types of accidents reported in the police records, i.e., fatal, serious, minor injury and property damage only. To get accident frequency, the data of all types of accidents were aggregated per year per section. This resulted in a dataset comprising a total of 222 samples out of which 148 randomly selected samples were used for training of models/algorithms and remaining 78 samples were used for validation. Based on the available data, 16 input variables were identified to predict the dependent variable, accident frequency (A). The summary statistics of these variables are provided in Table 2.

DNN model was built using DL4J (deep learning for Java) software, whereas a trial version of GeneXpro software [44] was used for GEP model. RENB model was built using SPSS software [45]. The models were validated using test data in terms of three different statistical measures, i.e., correlation coefficient (CC), root mean square error (RMSE) and mean absolute error (MAE).

2.2 Deep neural network (DNN)

The basic element of a back-propagation neural network (BPNN) is the processing node. Each processing node behaves like a biological neuron and performs two functions. First, it sums the values of its inputs. This sum is then passed through an activation function to generate an output. Any differentiable function can be used as an activation function, *f*. All the processing nodes in BPNN are arranged into layers, each fully interconnected to the following

Table 1 Dataset used for modeling of accident frequency

HID	Highway details		Sum of accidents								
	Highway name	Sections	2007	2008	2009	2010	2011	2012	2013	2014	Grand total
1	SH-11 (29.7 km)[a]	7				32	45	33			110
2	SH-20 (55.45 km)	9			41	32	33	46	34		186
3	SH-18(28.6 km)	8					39	35	43		117
4	MDR-137 (22.62 km)	7		77	88	78	68			68	379
5	NH-65 (23.38 km)	6	44	46	45	33					168
6	NH-73A (46.8 km)	7					59	38			97
7	NH-1 (79.35 km)	21				301	405	259	203		1168
8	NH-248A (22.5 km)	3		46	46	77	77	77	73	59	455
	Total	68	44	169	220	553	726	488	353	127	2680

[a]Length of highway covered is given in parenthesis

Table 2 Description of model variables for accident frequency modeling

Sr. no.	Variable with measurement units	Designation in the model	Min.	Max.	Mean	Std. deviation
1	Accident frequency (dependent variable)	A	0	96	12.07	15.780
2	Av. Daily Traffic (1000 PCU/day)	ADT	6.397	93.752	32.668	26.103
3	Section length (km)	L	0.400	13.200	4.891	3.314
4	Carriageway width (m)	RW	5.5	28.0	10.977	6.190
5	Paved shoulder width (m)	PSW	0.0	6.0	1.989	1.313
6	Median width (m)	MW	0.0	10.50	1.197	1.847
7	Number of minor access	MA	0	30	9.08	7.554
8	Number of horizontal curves	HC	0	19	3.84	4.493
9	Number of vertical curves	VC	0	6	0.423	0.867
10	Number of median openings	MO	0	15	2.34	4.023
11	Length of service road (km)	SR	0.000	18.100	1.070	3.203
12	Percentage of trucks in ADT	PT	11.34	56.19	34.47	12.61
13	Percentage of cars in ADT	PC	9.32	72.5	27.94	14.05
14	Driveways and commercial units along the road (numbers)	DW	0	54	9.41	9.932
15	No. of narrow bridges and culverts	BC	0	9	2.002	2.195
16	98th percentile speed (KMPH)	speed	66	123	80.52	12.149
17	Std. deviation of speed (KMPH)	STD	13.00	25.71	17.82	3.60

layer. There is no interconnection between the nodes of the same layer. In a BPNN, generally, there is an input layer that works as a distribution structure of the data being inputted to the network and not used for any type of processing. After this layer, one or more processing layers called the hidden layers follows. The final processing layer is called the output layer. A neural network with two or more hidden layers having large number of nodes and using sophisticated mathematical modeling is generally called deep neural network.

All the interconnections between each node have an associated weight. When a value is passed from the input layer, down these interconnections, these values are multiplied by the associated weight and summed to derive the net input (n_j) to the unit

$$n_j = \sum_i w_{ji} o_i$$

where w_{ij} is the weight of the interconnection to unit j from unit i (called input) and o_i is the output of the unit i. The net input obtained by the above equation is then transformed by the activation function to produce an output (o_j) for the unit j.

Traditionally, sigmoid and hyperbolic tangent are two extensively used nonlinear activation functions with a BPNN. Activation functions introduce nonlinearity in the neural network so as to learn more complex features present in the data. Saturation and sensitivity to changes around their mid-point were found to be two major problems with the sigmoid and hyperbolic tangent functions [46].

The rectified linear activation function (RELU; [47]) is a piecewise linear function and considered to be a major algorithmic change within last decades for the design of DNN [46]. RELU is one of the most popularly used activation functions in the deep learning which output the input value itself if it is positive, otherwise output would be zero. It is easier to train and found to achieve better performance than other activation functions with DNN. The RELU function is defined as:

$$f(n_j) = \max(0, n_j)$$

Random weight initialization is normally used with a standard BPNN because of the reason that stochastic gradient descent approach uses randomness in order to find optimal set of weights for the specific mapping function from inputs to outputs with the given dataset. Initializing BPNN with the correct weights is an important factor for proper functioning of neural network. The weights selected before the start of training the network should also be in a reasonable range.

Xavier weight initialization [48] was proposed as the weights initialization technique for DNN because of the poor performance of random weight initialization with standard gradient descent-based optimization approach. This approach assigns the weights from a Gaussian distribution with zero mean and some finite variance, thus allowing the variance of the outputs of a layer to be equal to the variance of its inputs.

Learning rate is also an important user-defined parameter used to adjust the weights of the BPNN. Most of the studies reporting the application of back-propagation neural network in civil engineering used learning rates values which were randomly set by the user (a value between 0 and 1) based on the past experiences and earlier reported works, suggesting it is often hard to get its correct values. Before the introduction of adaptive learning rate methods, the gradient descent algorithms used with BPNN were updating the network weights with the help of a learning rate, the objective function and its gradient. To improve the working of traditional gradient descent algorithms, adaptive gradient descent (using adaptive learning rate) algorithms, which could adaptively tune the learning throughout the training process, were proposed and used with DNN [46]. In this study, adaptive moment estimation (Adam; [49])-based optimization algorithm was used to update network weights during training in place of the classical stochastic gradient descent method. Adam computes individual learning rates for different parameters and requires setting several user-defined parameters. In this study, default values of all user-defined parameters as suggested by Kingma and Ba [49] were used and found working well with this data.

In the presence of limited training data, deep neural network may generally overfit (i.e., increase in generalization error), thus producing poor performance with test data. To avoid the problem of overfitting, Srivastava et al. [50] proposed the concept of introducing dropout layer in the design of DNN to improve their generalization capabilities. Except the choice of activation function, this layer requires a value called probability of retention in the hidden layers (i.e., p) to be defined by the user [50] which can be optimized through trial-and-error process.

Similar to simple BPNN, deep neural network required setting of several user-defined parameters. These parameters include, number and type of hidden layers, nodes in each hidden layer, activation function for output, hidden and dropout layers, weight initialization method, optimization algorithm, updaters (i.e., learning rate optimization algorithm), batch size (i.e., number of training samples used in one iteration) and number of epochs (i.e., one epoch is defined as when an entire training dataset has passed once through the neural network both in forward and backward directions).

2.3 Gene expression programming (GEP)

Proposed by Ferreira [28], gene expression programming (GEP) is a combination of both genetic algorithm (GA) and genetic programming (GP). GEP uses character linear chromosomes composed of genes structurally organized in a head and a tail. The chromosomes function as a genome and are subjected to modification by means of mutation, transposition, root transposition, gene transposition, gene recombination, and one- and two-point recombination. The chromosomes encode expression trees which are the object of selection. The creation of these separate entities (genome and expression tree) with distinct functions allows the algorithm to perform with higher efficiency that greatly surpasses existing adaptive techniques [28]. The genome is encoded as linear chromosomes of fixed length (just as in GA) that is then expressed as phenotype in the form of expression trees by GEP. Despite their fixed length, GEP chromosomes can code expression trees with different sizes and shapes [28]. The main aim of GEP is to find a mathematical function using a set of data.

In GEP, first of all a generation of initial population is created randomly that consists of individual chromosomes of fixed lengths. Each individual chromosome of the initial population is then evaluated by their fitness function which can be determined by user and then these chromosomes are allowed to reproduce with modifications by means of mutation, crossover and selection. This results in the development of new modified chromosomes of next generation which are again evaluated by fitness function. The process is repeated for a certain number of generations or until a good solution is found. For the mathematical equation, the GEP process performed the symbolic regression by means of the most of the genetic operators of GA.

2.4 RENB model

Regular negative binomial model (FENB model) accounts for over-dispersion but does not allow for spatial and temporal correlations found in the accident data collected repeatedly over interconnected roads. RENB model allows for random location and time effects by assuming that the over-dispersion parameter is randomly distributed across group. According to RENB model specification used in the study [51], the form of RENB model for the target μ (expected accident frequencies) with the random effects $\mathbf{\upsilon}$ can be defined as

$$\eta = \ln E(\mu|\mathbf{\upsilon}) = \mathbf{X}\beta + \mathbf{Z}\upsilon + \varepsilon, \quad \lambda|\mathbf{\upsilon} \sim NB, \tag{1}$$

where η is the linear predictor; ln is link function; $\mathbf{X}$ is a $(N \times p)$ design matrix for the p explanatory variables included in the model; $\boldsymbol{\beta}$ is a $p \times 1$ column vector of the fixed-effects regression coefficients; $\mathbf{Z}$ is a $(N \times r)$ design matrix for the r random effects (the random compliment to the fixed $\mathbf{X}$); $\mathbf{\upsilon}$ is a $(r \times 1)$ vector of random effects (the random compliment to the fixed $\boldsymbol{\beta}$) which are assumed to be normally distributed with mean 0 and variance matrix $\mathbf{G}$ (of random effects); $\boldsymbol{\varepsilon}$ is a $N \times 1$ column vector of the residuals, that part of μ which is not explained by the model; variance matrix of repeated measures is R and N is total number of observations. NB is the conditional target probability distribution. If there are no random effects, the model reduces to a regular or fixed effect negative binomial model.

Table 3 Performance indicators of various models

Fitness parameters	DNN model	GEP model	RENB model
RMSE	5.908	7.474	8.862
MAE	3.927	5.040	5.529
Correlation coefficient	0.945	0.914	0.891

3 Results

The comparative performance indicators of DNN, GEP and RENB model with test dataset are provided in Table 3. The results suggest that DNN model provided the best correlation coefficient in comparison with GEP and RENB model and resulted in the lowest root mean squared error. The scatter plot of observed and predicted values (Fig. 1) and line diagram (Fig. 2) showing comparison of predicted values by all models for test data also indicate superiority of DNN model over GEP and RENB model.

The user-defined parameters of the developed DNN and GEP models are provided in Table 4, whereas the final equation relating traffic accidents with different explanatory variables and variable importance provided by GEP model is provided in Table 5 and the expression trees of final GEP model, sub-ET 1 to 6 are shown in Fig. 3. The RENB model specifications and results are provided in Tables 6 and 7, respectively. The values provided in the parenthesis are standard errors of estimates.

4 Discussion

The results of this study clearly indicate that the performance of DNN model was best as compared to GEP and RENB model in terms of all the three performance indicators (Table 3). The observed v/s predicted accident frequency curves (Fig. 1 and 2) also underline predictive power of DNN model. Therefore, DNN model is most suitable model when the purpose of modeling is to predict accident frequency accurately to estimate related costs.

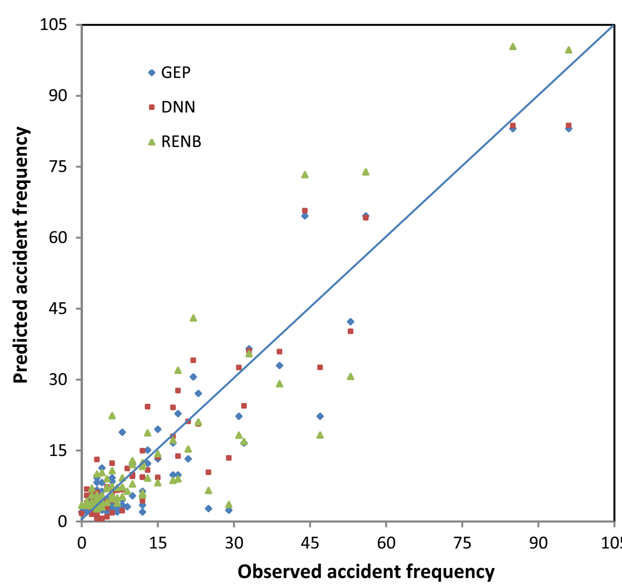

Fig. 1 Scatter plot of *observed versus predicted* accident frequency for test data

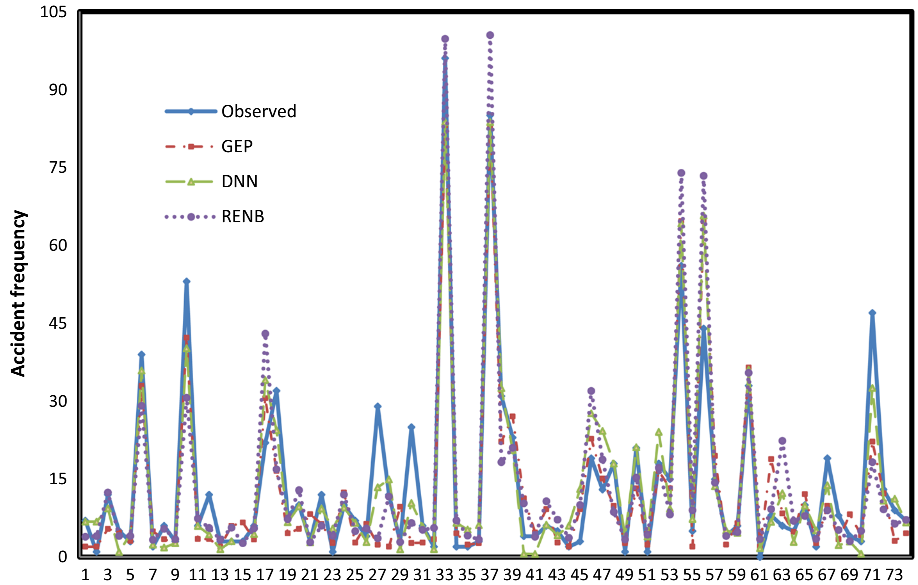

Fig. 2 Line diagram showing *observed versus predicted* accident frequency for test data

Table 4 Optimal value of user-defined parameters of DNN

Algorithm used	User-defined parameters	
DNN	Four hidden layers (200, 150, 100, 50 nodes), 5 dropout layer with $p = 0.5$, Activation function ReLU, weight initiation-XAVIER, batch size = 5, updater = Adam, epochs = 100	
GEP	Chromosomes	40
	Genes	6
	Head Size	12
	Functions used	Addition, multiplication, power and exponential
	Linking Function	Addition
	Fitness Function	RMSE

Table 5 GEP model equation and variable importance

Model equation	
A = 2.002 + MW*VC + 3*SR − 0.00624*VC*MO*L*BC + exp (−6.426)*DW*STD^2	
Variable	Importance
SR	0.545
DW	0.278
STD	0.109
VC	0.029
MW	0.026
MO	0.006
BC	0.004
L	0.003

GEP model was found performing better than RENB model (Table 3 and Figs. 1 and 2) with the used data. GEP model also provided a simple and easy-to-use equation (Table 5) to predict traffic accidents which can be used by field engineers.

Only five variables, namely length of service road, number of driveways, standard deviation of speed, number of vertical curves and median width, were considered most significant by the GEP model (Table 5). Accidents were found increasing with increases in service road length, number of driveways, median width, and standard

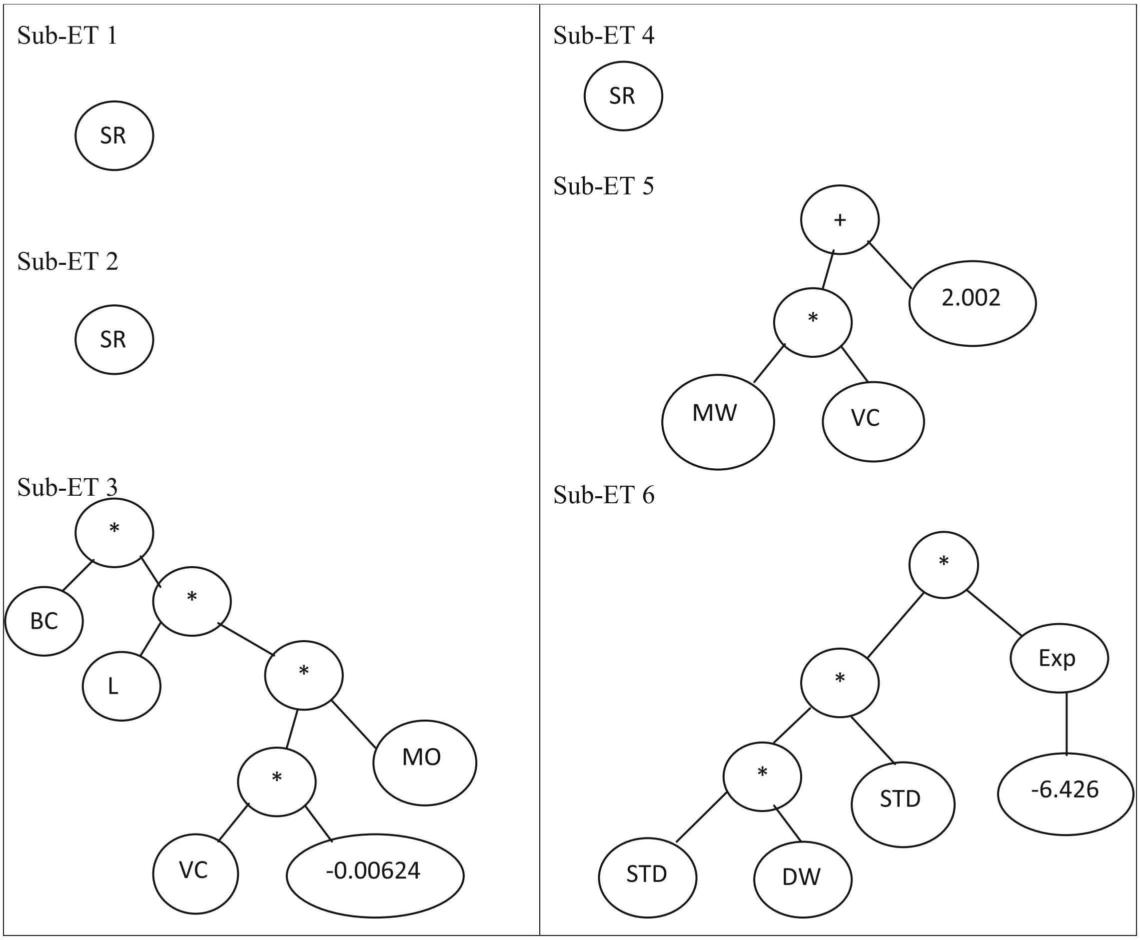

Fig. 3 The six GEP expression trees for developing accident prediction model

Table 6 Model specifications for random effect model

Sr. No.	Description	Effects	RENB model with covariance structure AR1
			Number
1	Covariance parameters	Residual effects	2
		Random effects	1
2	Design matrix	Fixed effects	7
		Random effects	1[a]
3	Common subjects		63

Common subjects are based on subject specifications for random and residual effects and are used to chunk the data for better performance. Subject specification: HID*RID

[a]This is the number of columns per common subject

deviation of speed and decreasing with increasing length of road section, median openings, number of narrow bridges and culverts. GEP model included interaction terms also. Increase in median width when associated with vertical curves results in higher number of accidents. Similarly, standard deviation of speed which signifies mixed nature of traffic, when associated with more number of driveways, results in higher accidents. In the same way, increase in the number of vertical curves when associated with increase in median opening and number of narrow bridges and culverts results in reduction in accidents. On the other hand, RENB model (Tables 6, 7) suggests six input variables having significant effect on accidents: traffic volume, length of section, paved shoulder width, number of minor accesses, length of service road and standard deviation of speed. Out of these six significant variables, only paved shoulder width was found negatively associated with accident frequencies. RENB model results indicate that keeping all

Table 7 Parameter estimates of RENB model equation

Sr. No.	Model characteristics	RENB model with covariance structure auto-regressive-1 (AR-1)
A	*Model statistics*	
1	Bayesian information criterion (BIC)	288.740
2	Akaike information criterion-finite sample corrected (AICC)	280.092
3	− 2 log pseudo likelihood	273.915
B	*Parameter estimates*	
1	(Intercept)	− 0.525 (0.394)[a]
2	ln ADT	0.217 (0.098)
3	ln L	0.148 (0.071)
4	M_Acc	0.038 (0.007)
5	PSW	− 0.143 (0.044)
6	SR	0.105 (0.010)
7	STD	0.083 (0.018)
	Over-dispersion parameter	0.174
C	*Covariance parameters*	
a.	Residual effect	
1	AR1 Diagonal	0.605 (0.109)
2	AR1 Rho	− 0.004 (0.187)
b.	Random effects	0.152(0.052)
D	*Model equation*	$A=0.592*ADT^{0.217}*L^{0.148}*e^{(0.038*M\_Acc+0.105*SR+0.083*STD-0.143*PSW)}$

[a]Values in parenthesis are estimates of standard error

other effects constant, every one meter increase in paved shoulder width reduces accidents by 13.3%. On the other hand, increase in the number of minor access points by one increases accidents by 3.9%, increase in standard deviation by 1 kmph increases accidents by 8.7%, every 1 km increases in service road length increases accidents by 11.1% and increment of 1 in value of natural log of traffic volume increases accidents by 24%.

Both, GEP and RENB models consider service road as the most significant variable having positive association with accidents, which in the present study can be a result of dysfunctional, ill-designed and ill-maintained service roads with no control of access in the study area [26]. Accident frequency was found to increase with increase in exposure in terms of traffic volume and the distance travelled as suggested by RENB model. This finding is in agreement with the results reported in the literature [24, 52]. Both models considered standard deviation of speed, which reflects the mixed nature of traffic, as significant variable positively contributing to accidents. These results are in accordance with studies reported in the literature [53, 54]. GEP model provided further insights by suggesting that the contribution of mixed traffic toward accidents is more when the highway has more number of driveways or road side development. Median width was found to be positively associated with the accident frequencies as indicated by GEP model. Median width increases the number of all accident types, possibly due to the difficulty during over-taking maneuvers, presence of a new hazard in the car-riageway [55] and due to its association with increased speeds [56], particularly during night time. Speeds are higher on divided roads and still higher on wider median road sections due to the absence of glare during night.

Similar to the study reported by Ackaah and Salifu [57], minor accesses on highways were found positively associated with accidents as demonstrated by RENB model, but GEP model considered number of driveways as significant variable in the place of number of minor accesses. Positive association of driveway density with accidents has also been reported by various researchers [24, 58, 59]. GEP model further suggests that combined effects of increased number of median openings, bridges and culverts and vertical curves cause reduction in the accidents. This seems reasonable as driver becomes more cautious on such stretches [13].

5 Conclusion

This paper investigates the potential of DNN and GEP model by comparing their results with conventional RENB regression model for predicting the road accident frequencies. A major conclusion from this study is that DNN model outperformed both GEP and RENB model on all performance indicators, but it provided no quantification of effects of various causal variables on accidents. Therefore, it can be used when the purpose of modeling is to predict accident frequency accurately to estimate related costs. GEP model also performed better than RENB model, and it also has the capability of identifying important risk factors and quantifying their effect on accidents, suggesting it as an effective alternate to RENB model for traffic accident prediction when the purpose of modeling is to suggest improvement measures based upon some logical findings.

Based on the discussion presented in Sect. 4 the following recommendations can be made:

1. On highways, service roads are required to be properly designed with smooth and gradual exit and entry and made functional by removing encroachment.
2. Direct entry of minor access roads into highway must be restricted by providing service roads with well-designed entry and exit points, and there should also be a check on road side development to restrict the growth of number of driveways and commercial establishments along the road.
3. The speed-control strategies and imposition of speed limits are urgently required on highways to make the traffic mix more uniform.
4. Provision of paved shoulder is recommended, especially on roads with large number of minor access points, carrying mixed traffic as was the case in the present study.

References

1. MORTH (2017) Basic road statistics, 2017. Ministry of Road Transport and Highways, Government of India, India
2. NIC (2017) State-wise distribution of accidental deaths by unnatural causes during 2017. National Informatics Centre Department of Electronics and Information Technology, Ministry of Communications and Information Technology, Government of India
3. MORTH (2018) Road accidents in India, 2018. Ministry of Road Transport and Highways, Government of India, India
4. Shaheem S, Mohammed KMS, Rajeevan (2006) Evaluation of cost effectiveness of improvements of accident prone locations on NH-47 in Kerala state. Indian Highw 34(2006):35–46
5. Elvik R (2011) Assessing causality in multivariate accident models. Accid Anal Prev 43:253–264
6. Savolainen PT, Mannering FL, Lord D, Quddus MA (2011) The statistical analysis of highway crash-injury severities: a review and assessment of methodological alternatives. Accid Anal Prev 43(5):1666–1676
7. Hauer E (2010) On prediction in road safety. Saf Sci. https://doi.org/10.1016/j.ssci.2010.03.003
8. Lord D, Mannering F (2010) The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives. Transp Res Part A 44:291–305
9. Mohan D, Tsimhoni O, Sivak M, Flannagan, MJ (2009) Road safety in India: challenges and opportunities. UMTRI, 1
10. Mohan D (2002) Traffic safety and health in Indian Cities. J Transp Infrastruct 9(1):79–94
11. MORTH (2007) Sundar committee report, February, 2007. Ministry of Road Transport and Highways
12. Sharma AK, Landge VS, Deshpande NV (2013) Modeling motorcycle accidents on rural highway. Int J Chem Environ Biol Sci 1(2):313–317
13. Jacob A, Anjaneyulu MVLR (2013) Development of crash prediction models for two-lane rural highways using regression analysis. Highw Res J 6(1):59–70
14. Fletcher JP, Baguley CJ, Sexton B, Done S (2006) Road accident modeling for highway development and management in developing countries. Main Report: Trials in India and Tanzania. Project Report No: PPR095, DFID
15. Lord D, Geedipally SR, Persaud BN, Washington SP, Ivan JN, van Schalkwyk I, Lyon C, Jonsson T (2008) Methodology for estimating the safety performance of multilane rural highways. NCHRP Web-Only Document 126, National Cooperation Highway Research Program, Washington, DC. http://onlinepubs.trb.org/onlinepubs/nchrp/nchrpw126.pdf. Accessed 3 June 2010
16. Cafiso S, Graziano AD, Silvestro GD, Cava GL, Persaud B (2010) Development of comprehensive accident models for two lane rural highways using exposure, geometry, consistency and context variables. Accid Anal Prev 42:1072–1079
17. Hosseinpour M, Yahaya AS, Sadullah AF (2014) Exploring the effects of roadway characteristics on the frequency and severity of head-on crashes: case studies from Malaysian Federal. Accid Anal Prev 62:209–222
18. Caliendo C, Guglielmo MLD, Guida M (2015) Comparison and analysis of road tunnel traffic accident frequencies and rates using random-parameter models. J Transp Saf Secur 8(2):177–195
19. Kibar FT, Celik F, Wegman F (2017) Analyzing truck accident data on the interurban road Ankara–Aksaray–Eregli in Turkey: comparing the performances of negative binomial regression and the artificial neural networks models. J Transp Saf Secur. https://doi.org/10.1080/19439962.2017.1363841
20. Abusini S, Kilicman A (2017) Modeling motorcycle accident on the road section by using general linear model: case studies in Batu City. Cogent Math 4(1):1379675. https://doi.org/10.1080/23311835.2017.1379675
21. Gianfranco F, Soddu S, Fadda P (2018) An accident prediction model for urban road networks. J Transp Saf Secur 10(4):387–405. https://doi.org/10.1080/19439962.2016.1268659
22. Xie Y, Lord D, Zhang Y (2007) Predicting motor vehicle collisions using Bayesian Neural Network models: an empirical analysis. Accid Anal Prev 39:922–933
23. Driss M, Benabdeli K, Hamadouche T, Saint-Gerand MA (2015) Traffic safety prediction model for identifying spatial degrees of exposure to the risk of road accidents based on fuzzy logic approach. Geocarto Int 30(3):243–257. https://doi.org/10.1080/10106049.2014.883554
24. Singh G, Sachdeva SN, Pal M (2018) Comparison of three parametric and machine learning approaches for modeling accident severity on non-urban sections of Indian highways. Adv Transp Stud Int J Sect B 45:123–140
25. Chang LY, Wang HW (2006) Analysis of traffic injury severity: an application of non-parametric classification tree techniques. Accid Anal Prev 38(5):1019–1027
26. Singh G, Sachdeva SN, Pal M (2016) M5 model tree based predictive modeling of road accidents on non-urban sections of highways in India. Accid Anal Prev 96:108–117

27. Schmidhuber J (2015) Deep learning in neural networks: an overview. Neural Netw 61:85–117
28. Ferreira C (2001) Gene expression programming: a new adaptive algorithm for solving problems. Complex Syst 13(2):87–129
29. Zhang Z, Wang Y, Chen P, Yu G (2018) Application of long short-term memory neural network for multi-step travel time forecasting on urban expressways. In: CICTP 2017: transportation reform and change—equity, inclusiveness, sharing, and innovation. American Society of Civil Engineers, Reston, pp 444–454
30. Zhou L, Chen X (2018) Short-term forecasting of traffic flow and speed: a deep learning approach. In: CICTP 2018: intelligence, connectivity, and mobility. American Society of Civil Engineers, Reston, VA, pp 2186–2196
31. Deng F, He Y, Zhou S, Yu Y, Cheng H, Wu X (2018) Compressive strength prediction of recycled concrete based on deep learning. Constr Build Mater 175:562–569
32. Kumar SS, Abraham DM (2019) A deep learning based automated structural defect detection system for sewer pipelines. In: Computing in civil engineering 2019: smart cities, sustainability, and resilience. American Society of Civil Engineers, Reston, VA, pp 226–233
33. Dick K, Russell L, SouleyDosso Y, Kwamena F, Green JR (2019) Deep learning for critical infrastructure resilience. J Infrastruct Syst 25(2):05019003
34. Ding F, Zhang Z, Zhou Y, Chen X, Ran B (2019) Large-scale full-coverage traffic speed estimation under extreme traffic conditions using a big data and deep learning approach: case study in China. J Transp Eng Part A Syst 145(5):05019001
35. Nguyen T, Kashani A, Ngo T, Bordas S (2019) Deep neural network with high-order neuron for the prediction of foamed concrete strength. Comput Aided Civil Infrastruct Eng 34(4):316–332
36. Yu P, Yan X (2019) Stock price prediction based on deep neural networks. Neural Comput Appl. https://doi.org/10.1007/s00521-019-04212-x
37. Zhou Y, Zhang L, Yi Z (2019) Predicting movie box-office revenues using deep neural networks. Neural Comput Appl 31:1855. https://doi.org/10.1007/s00521-017-3162-x
38. Lecun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436
39. Nazari A (2019) Predicting the total specific pore volume of geopolymers produced from waste ashes by gene expression programming. Neural Comput Appl 31(2):751. https://doi.org/10.1007/s00521-012-1135-7
40. Gholami A, Bonakdari H, Zeynoddin M et al (2019) Reliable method of determining stable threshold channel shape using experimental and gene expression programming techniques. Neural Comput Appl 31:5799. https://doi.org/10.1007/s00521-018-3411-7
41. Malik H, Mishra S (2016) Application of gene expression programming (GEP) in power transformers fault diagnosis using DGA. IEEE Trans Ind Appl 52(6):4556–4565
42. Dindarloo SR, Pollard JP, Siami-Irdemoosa E (2016) Off-road truck-related accidents in US mines. J Saf Res 58:79–87
43. Yaacob W, Lazim M, Wah Y (2010) Evaluating spatial and temporal effects of accidents likelihood using random effects panel count model. In: Proceedings of the 2010 international conference on science and social research, Kuala Lumpur, Malaysia
44. GENexpro (5.0 trial version) https://www.gepsoft.com/. Accessed 21 Feb 2018
45. IBM Corp. Released 2013. IBM SPSS Statistics for Windows, Version 22.0. IBM Corp., Armonk, NY
46. Goodfellow I, Bengio Y, Courville A (2016) Deep learning. MIT Press, USA
47. Nair V, Hinton GE (2010) Rectified linear units improve restricted boltzmann machines. In: Proceedings of the 27th international conference on machine learning (ICML-10), Haifa, Israel, June 21–24, pp 807–814
48. Glorot X, Bengio Y (2010) Understanding the difficulty of training deep feedforward neural networks. In Proceedings of the thirteenth international conference on artificial intelligence and statistics, pp 249–256
49. Kingma DP, Ba J (2015) Adam: a method for stochastic optimization. In: 3rd international conference on learning representations, May 7–9, 2015, San Diego, https://arxiv.org/abs/1412.6980. Accessed 15 June 2019
50. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R (2014) Dropout: a simple way to prevent neural networks from overfitting. J Mach Learn Res 15(1):1929–1958
51. Agresti A, Booth JG, Caffo B (2000) Random-effects modeling of categorical response data. Sociol Methodol 30:27–80
52. Dinu RR, Veeraragavan A (2011) Random parameter models for accident prediction on two-lane undivided highways in India. J Saf Res 42(1):39–42
53. Quimby A, Maycock G, Palmer C, Buttress S (1999) The factors the influence a driver's choice of speed: a questionnaire study. Transport Research Laboratory, Crowthorne, UK
54. Taylor MC, Baruya A, Kennedy JV (2002) The relationship between speed and accidents on rural single-carriageway roads, vol 511. TRL, Crowthorne
55. Elvik R, Vaa T (2004) The handbook of road safety measures. Elsevier, Amsterdam
56. Hauer E (2000) The median and safety. www.trafficsafetyresearch.com (unpublished)
57. Ackaah W, Salifu M (2011) Crash prediction model for two-lane rural highways in the Ashanti region of Ghana. IATSS Res 35(1):34–40
58. Harwood DW, Council FM, Hauer E, Hughes WE, Vogt A (2000) Prediction of the expected safety performance of rural two-lane highways (No. FHWA-RD-99-207, MRI 4584-09, Technical Report). Federal Highway Administration, United States
59. Fitzpatrick K, Lord D, Park BJ (2010) Horizontal curve accident modification factor with consideration of driveway density on rural four-lane highways in Texas. J Transp Eng 136(9):827–835