ESTIMATION OF VULNERABLE ROAD USER ACCIDENT FREQUENCY THROUGH THE SOFT COMPUTING MODELS

Saurabh Jaglan[1,*], Sunita Kumari[1], Praveen Aggarwal[2]

[1]Civil Engineering Department, Deenbandhu Chhotu Ram University of Science and Technology, Murthal, Sonepat, Haryana, India
[2]Civil Engineering Department, National Institute of Technology, Kurukshetra, Haryana, India

*E-mail of corresponding author: saurabhjaglan.civil@dcrustm.org

Saurabh Jaglan 0000-0001-8719-4677, Sunita Kumari 0000-0003-3050-5096

Resume

Accident prediction models are mathematical expressions or algorithms used to determine the causal factors for road accidents and road safety engineers are using these models, as well. Modelling this kind of accident is quite challenging and required good quality of data. The results of the artificial neural network model, Gaussian processes model, and support vector machine model are compared for vulnerable road accident frequency in this study. The accident frequency dataset comprises 218 records, with 146 designated for training purposes and 72 reserved for testing. The model's accuracy was contingent on: the mean absolute error, root mean square error and coefficient of correlation. The findings suggest that for predicting the vulnerable road user accidents on roads, the artificial neural network gives better correlation results as (0.912) that the support vector machine (0.879) and Gaussian processes (0.853).

Article info

Received 27 October 2023
Accepted 21 February 2024
Online 11 March 2024

Keywords:

vulnerable road user
accident frequency
artificial neural network
support vector machine and
Gaussian processes

Available online: https://doi.org/10.26552/com.C.2024.023

ISSN 1335-4205 (print version)
ISSN 2585-7878 (online version)

1 Introduction

Almost 400 million trucks and 1.1 billion cars were driven globally in 2015. It is anticipated that 800 million trucks and 2 billion cars would be driven worldwide in 2040 [1]. The increase in vehicular population causes an increase in accidents worldwide, as well as in India. According to Road accidents in India, there has been an alarming annual increase in accidents of 11.9% and a rise in fatalities of 9.4%. These road safety issues in India have been a major concern, and accidents involving vehicles and vulnerable road users (VRUs) are unfortunately common [2]. The VRUs include pedestrians, cyclists, and two-wheeler riders. Several factors, including road characteristics, traffic-related variables, environmental factors, speed characteristics, etc., influence *the* vulnerable road user (VRU) accidents. Since the VRU road accident data is discrete and non-negative, it mostly counts data models used in accident frequency modelling (i.e., the number of accidents on a specified section in known time duration). Various data sets were collected from roadside geometry and first information report (FIR) data collection to estimate the accident frequency. The random division method divides the entire dataset into two parts per the study's requirement. For this study data set was collected on about 364 kms of road length, including different national highway and state highways of Haryana a northern state of India. The roads are then divided into sub-sections based on other geometric data. Two hundred eighteen accident frequency records were divided into two parts. Out of these, 146 samples were used to create the model (training set), and 72 samples were used to test the model in the study. Traditional methodologies are employed using *the* assumed analytical equations to evaluate the impact of various road geometries and accident-related factors, and conducting regression. Nonetheless, those methods prove challenging to apply, and the existing literature offers no completely accurate predictions for vulnerable road user accidents. The flowchart of this investigation is shown in Figure 1. Figure 2 represents the online portal of the Haryana police accident recording application. This portal is used to extract the VRU accident-related information. This portal requires the year, district and police station as input data and then provides the complete details of every accident.

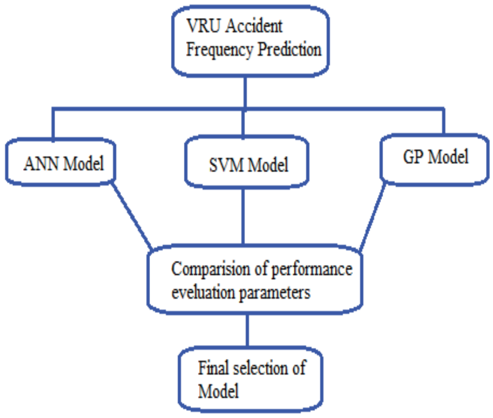

Figure 1 *The process of investigation*

Figure 2 *Haryana Police Online Portal App*

2 Literature review

Various artificial intelligence techniques, including Gaussian processes (GP), support vector machines (SVM), artificial neural networks (ANN), random forest (RF), adaptive network-based fuzzy inference system (ANFIS), M5P model trees, have gained significant popularity and widespread utilization among researchers in different fields [3-5]. Many of these studies suggest that artificial intelligence techniques exhibit considerably high accuracy. Numerous researchers employed the M5P model tree, artificial neural network, random forest model [6-8], with these approaches yielding the best-suited results. This research aims to predict the total VRU Accident frequency by comparing it using GP, SVM, and ANN methods. The recent scenario suggests that the frequency of vulnerable road accidents depends on many variables. Various researchers considered the road parameters and accident characteristics to impact accident frequency [9-11]. Those parameters may be *the* road width, shoulder width, length of section, traffic, median opening, service road, commercial units, etc. Nowadays, different studies are conducted to evaluate the effect of other variables length (L), Shoulder width (SW), road width (RW), median opening (MO), median assess (MA), average daily traffic (ADT), service road (SR) etc. on accidents and the impact of accidents [12-16]. In this study *the* SVM, ANN and GP models were applied using Waikato Environment for Knowledge Analysis, version- 3.9.5 software. The present study meticulously applies various advanced soft computing techniques to address the critical issue of predicting vulnerable road user accident frequency (VRUAF). These innovative and adaptive methodologies, rooted in artificial intelligence and data analysis, offer a dynamic and flexible approach to tackling the complex challenges of road safety for VRUs. The diverse set of soft computing techniques employed in past research includes [15].

2.1 Artificial neural network

This study uses an Artificial Neural Networks (ANN) Model to predict *the* accident frequency involving vulnerable road users. The architecture of an ANN, designed for vulnerable road user accident severity prediction; typically involves multiple layers of interconnected neurons is used [17-18]. The input layer receives relevant features, such as average daily traffic (ADT), Length of section, shoulder width, vehicle

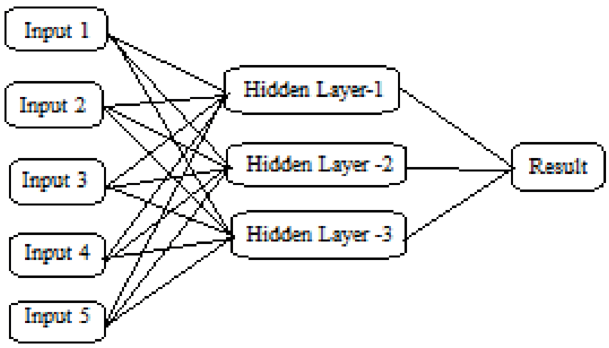

Figure 3 *Nomenclature of ANN model*

speed, time of day, and more. Hidden layers, which can be customized in depth and width, process these inputs to learn intricate relationships and patterns within the data. These layers of the ANN are responsible for processing the input data through a series of weighted connections and non-linear activation functions. Different models were generated through experimentation and optimization to achieve the best predictive performance by changing neurons and the number of hidden layer neurons. Various ANN models were generated with different hidden layers and training times to get better results. Neurons in the hidden layers apply mathematical transformations to the input data to capture complex combinations of input variables. The output layer of the ANN provides the predicted output, which in this study is the estimated frequency of VRU accidents. Typically, there is a single output node for the regression tasks like accident frequency prediction. The output value represents the predicted accident frequency. Following the establishment of the optimal architecture, the dataset comprising a total of 218 accidents should be segregated into two distinct subsets. The first set is for training the network, and the second is exclusively reserved for testing the network's performance. In this study the complete accident data is divided into two parts, i.e., *the* training set and testing set. Firstly, using 70% of the total data, the model is trained using all the parameters for various correlation coefficient (CC) values. The same trial is to be performed many times by changing the neurons in hidden layers. The remaining 30% of the data is utilized to validate the developed model in terms of the best-suited CC value. Figure 3 represents the working of the ANN model with different input variables and hidden layers. Various combinations are formed with different hidden layers, and the best suited model is considered as the result of the ANN model.

2.2 Support vector machines

Regression using support vectors and classification techniques from statistical learning theory are known as support vector machines. The ideal separation of classes is the foundation of classification techniques based on the support vector machines. This approach chooses the linear classifier that minimizes the generalization error, or at least an upper bound on this error, resulting from structural risk minimization, from an infinite pool of potential classifiers, provided that the classes can be distinguished from one another. As a result, the hyperplane that leaves the greatest margin between the two classes will be chosen; according to Cortes and Vapnik (1995), the margin is the total of the hyperplane's distances from the points closest to the two classes [19].

2.3 Gaussian process

One easy to understand and popular family of probability distributions on functions is the Gaussian process (GP). Viewed in this broad context, scholars have researched and employed many varieties of Gaussian processes. However, this focuses on the more focused use of Gaussian processes to predict. Distributions over functions are represented by a Gaussian process using training data and testing data. We have a training data, T= {A,b}, where, A= $\{a_1, a_2, a_3, \ldots a_n\}$ matrix have x_i as input examples, and b = $[b_1, b_2, b_3, b_4 \ldots b_n]$ vector have b_j as training output. Training data T = {A,b} with the input test as z* [20]. Where the input vectors of test set matrix are, $z^* = [z_1, \ldots z_n]$

The following equation gives the overall observations:

$$P_i = f(z_i) + \varepsilon, \quad (1)$$

where the P_i is used to predict the results using testing data set. Function $f(z_i)$ is the function/parameters used to prediction and variance and mean zero is taken as ε in the equation.

2.4 Details of Kernel functions

The kernel function concept was proposed on data to simplify it into a higher dimensional feature involving non-linear decision surfaces. Without actually doing

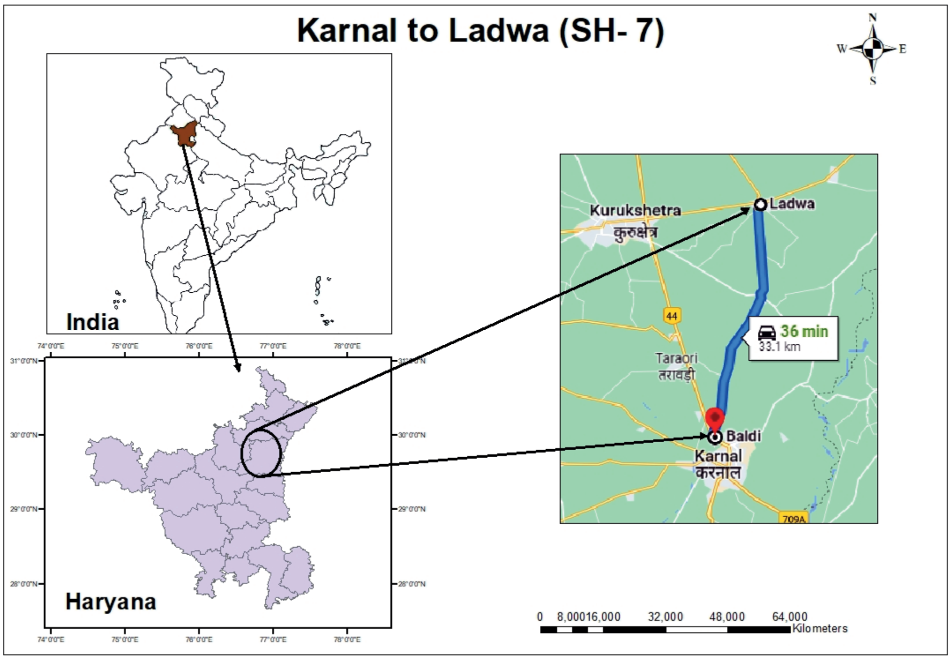

Figure 4 *One of the Study areas for Accident and road geometry data collection*

Table 1 *Description of model variables for accident frequency and severity modelling*

	Descriptive Statistics				
	N	Minimum	Maximum	Mean	Std. Deviation
ADT	218	4127	105673	26913.92	30558.705
Ln ADT	218	8.325	11.568	9.7678	0.853
L	218	.7	19.0	5.442	3.589
Ln L	218	-0.356	2.944	1.455	0.745
RW	218	5.5	28.0	13.115	7.062
SW	218	0	6.0	2.349	1.741
Median Width	218	0	4.50	.6709	1.017
Minor Access	218	0	24	8.47	5.989
HC	218	0	18	4.96	4.784
VC	218	0	6	1.44	1.505
MO	218	0	11	4.39	2.996
SR	218	0	4.50	1.406	1.133
CAR %	218	31.330	73.290	49.608	12.000
TRUCK/HV %	218	11.23	44.68	19.837	8.563
Commercial Units	218	0	47	11.57	11.539
B and C	218	0	22	3.57	4.849
SPEED	218	20	110	65.47	21.147
A/Yr	218	1	32	10.26	6.981

Where L = Length of section (km); LnL = Log of Length of section; ADT = Average Daily Traffic; LnADT = Log of Average Daily Traffic; SW = Shoulder Width; RW = Road Width; MO = Median Opening; MA = Minor Access; MW = Median Width; HC = Horizontal Curve; VC = Vertical Curve;

C % = Percentage of cars; T % = Percentage of Truck; B & C = Bridge and Culverts; A/yr = Accidents per year; SR = Service Road.

calculations in a high-dimensional feature space, any selected model can operate there because of a kernel function. The selection of appropriate kernel parameters is crucial and often involves hyperparameter tuning to optimize the model performance. A kernel function provides a way to measure the similarity or dissimilarity between the data points and is a key component in many machine learning models, enabling them to operate effectively in higher-dimensional spaces without explicitly computing the transformed feature vectors. A mathematical function is a kernel function, and any symmetric positive semidefinite function that meets Mercer's requirements can be utilized as a kernel function, according to Cortes and Vapnik (1995) [19]. Several kernel functions are available, but selecting the one that provides the optimum generalization with a particular dataset might be challenging. The selection of kernel should be as per requirement for the underlying structure of the data. Cross-validation and other model selection techniques are often employed to determine the best kernel for a specific problem. The correlation coefficients (CC) and Root Mean Square Error (RMSE) were examined to determine the best selection of these factors. Different models were created using the input dataset as a training set to forecast the VRU accident frequency. The regression model's accuracy is estimated through testing. In the current study, the effectiveness of different models in predicting the efficiency of the VRU accident frequency was evaluated using the correlation coefficient, R2, and the root mean square error (RMSE).

3 Method and materials

Data were derived from several resources, including the Harsamay portal of Haryana Police, field visits, the PWD department, toll plaza, etc. Data was collected for accidents and road-related elements, as shown in Figure 4. The range of accident frequency lies between 0 and 32 at the single road starch. The complete VRUAF dataset comprises 218 data sets divided into training and testing set. The descriptive statics of input and output variables are shown in Table 1. The minimum, maximum, mean and standard deviations are given for each parameter used in the study.

4 Result and discussion

Different models were developed using the collected VRU accident data set to predict the accidents with other influencing parameters. The first and most efficient ANN model is applied by changing the hidden layers from 1 to 5. The different CC, RMSE, and MAE values are calculated for each model approach. The best-suited highest value is taken as the final for three numbers of hidden layers.

4.1 Support vector machine model results

The training and testing sets included 70% and 30% of the entire data set, respectively. Several manual trials were performed for the SVM using different Kernel Function based model development. The results in the form of performance evaluation parameters (CC, RMSE, and MAE) are listed in Table 2 for both the training and testing stages. Different models were generated with the help of a training data set of vulnerable road use accidents. After the model development, the same model was applied to the testing data set for evaluating their performance. The conclusions provided in Table 2 show that the Pearson Vii kernel function based SVM model exhibits better results than all the other applicable models. Specifically, the training stage values of CC, RMSE, and MAE are 0.9064, 0.0820, and 0.2586, while the testing stage values are 0.8795, 1.5630, and 2.3681, respectively. The agreement plot between the actual and predicted VRU accident frequency values using the SVM_puk model for the testing stage is shown in Figure 5. All predicted values lie closer to the line of perfect agreement (1:1) line.

Figure 5 shows the plots illustrating the actual VRUs accident frequency distribution and predicted values of the VRUs accident frequency using the SVM_ puk model. The representation of this figure evidently demonstrates that the SVM_puk-based model predicted

Table 2 *SVM model results using different Kernel Function*

S.No	Kernel Function	SVM					
		TRAINING SET (146)			TESTING SET (72)		
		Correlation Coefficients	Mean Absolute Error	Root Mean Square Error	Correlation Coefficients	Mean Absolute Error	Root Mean Square Error
1	SVM_ RBFKernal	0.8728	1.1214	1.5581	0.8589	1.4434	1.9114
2	SVM _Puk	0.9064	0.0820	0.2586	0.8795	1.5630	2.3681
3	SVM_ PolyKernel	0.8475	1.3485	1.8149	0.8326	1.5821	2.4245

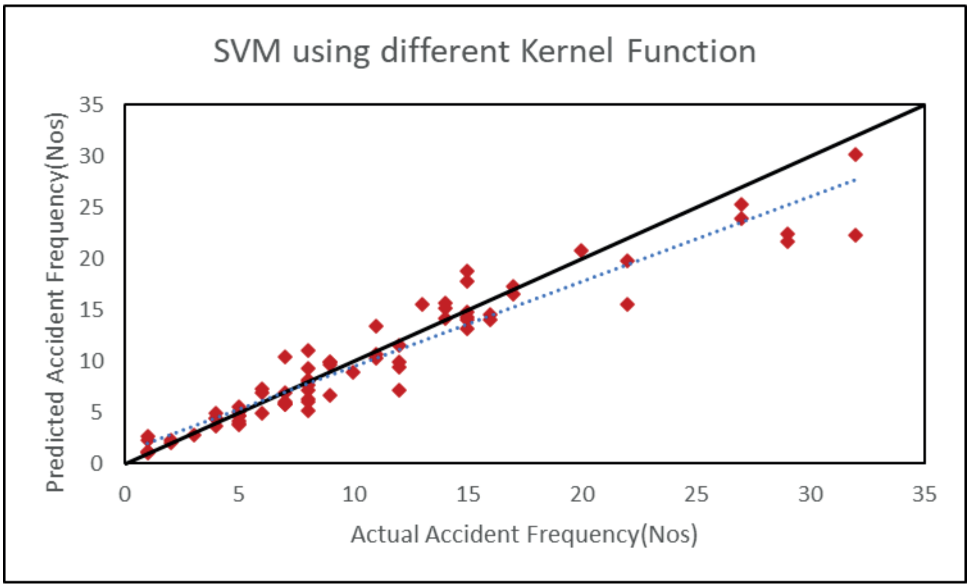

Figure 5 *Variation in scatter values with the actual values of the VRUAFs using SVM Model*

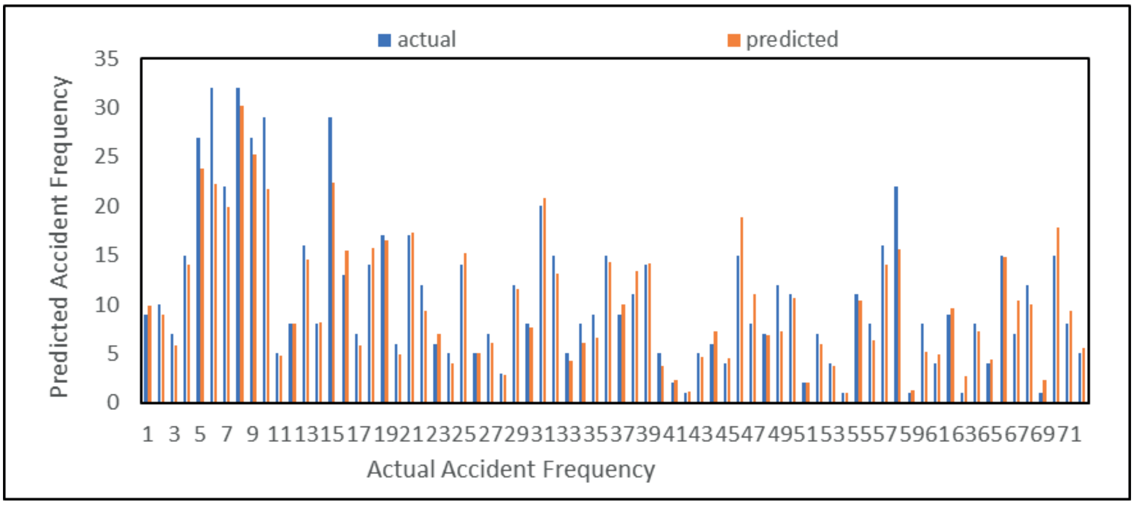

Figure 6 *SVM Model variation in actual and predicted VRUAFs values*

closer values to actual VRUAFs' values in comparison to the other models. The VRUs accident frequency using the SVM Model with actual and predicted values are shown in Figure 6.

4.2 Artificial neural network model results

In this study, the architecture of an ANN designed, for vulnerable road user accident frequency prediction, typically involves multiple layers of interconnected neurons.

The table 3 indicated the results with different hidden layers keeping the training time constant using ANN Model approach. The training time was taken as 1000 for each model formulation. Various models were generated using the value 0.2 of learning rate and 0.1 of momentum. Then, the same model was applied to the testing data set, and the results are given in Table 3. The same procedure was repeated for all five ANN models by changing the hidden layers, and results were calculated in terms of CC, RMSE, and MAE. The best value of ANN-4 is considered in terms of CC, RMSE, and MAE values are 0.9245, 1.458, and 1.123, respectively for the training set. The final ANN-4 model was used for further comparison with other modelling techniques in terms of these three performance indicators for VRUs accident frequency data. The scatter diagram of the actual and predicted VRU accident frequency values using the ANN-4 model in the testing dataset is shown in Figure 7. Figure 8 represents the predicted and actual VRUAF values with the help of the ANN Model.

4.3 Gaussian processes model result

Model development of GP is similar to SVM based model development. The results for training and testing data sets of various kernel function based GP models in for CC, RMSE, and MAE are listed in Table 4. The

Table 3 *Results of neurons-based ANN models*

ANN Model								
S. No	Model	No of neurons in hidden Layer	TRAINING SET (146)			TESTING SET (72)		
			Correlation Coefficients	Mean Absolute Error	Root Mean Square Error	Correlation Coefficients	Mean Absolute Error	Root Mean Square Error
1	ANN-1	1	0.834	1.183	1.516	0.802	1.245	1.987
2	ANN-2	2	0.878	1.041	1.383	0.852	1.396	1.799
3	ANN-3	3	0.912	0.947	1.252	0.904	1.337	1.741
4	ANN-4	4	0.924	1.124	1.458	0.912	1.198	1.799
5	ANN-5	5	0.895	1.585	1.651	0.860	1.685	1.854

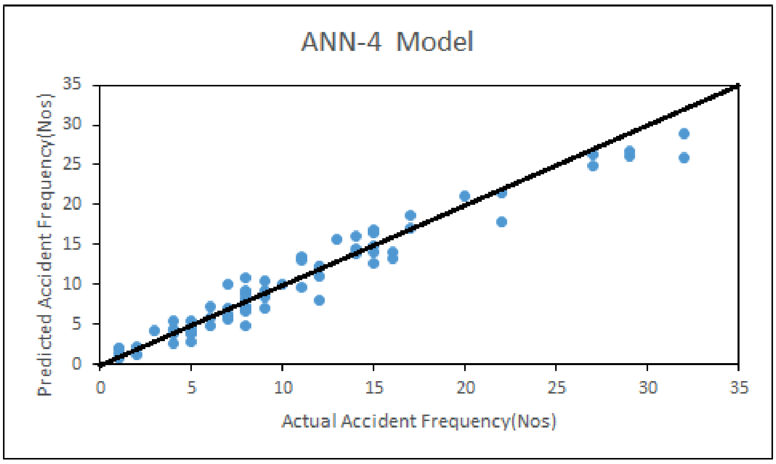

Figure 7 *ANN-4 model results for predicted and actual values of VRUAF*

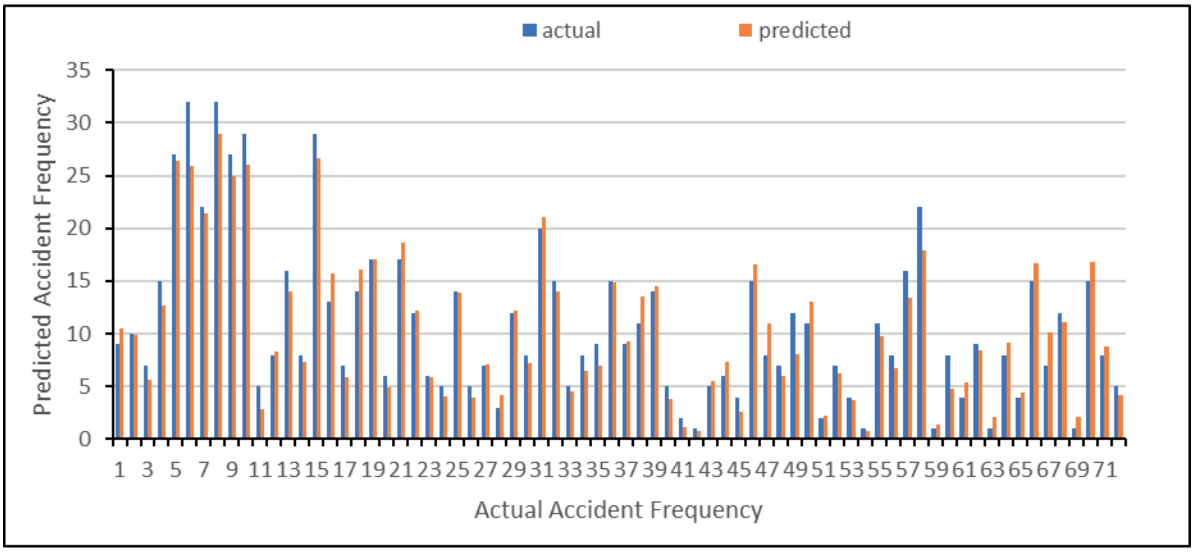

Figure 8 *ANN model comparison of predicted and actual values of VRUAF*

results of Table 4 demonstrate that the Pearson Vii kernel function-based GP model performs the best out of all the applied models. Its CC, RMSE, and MAE values are 0.885, 1.231, respectively, and 1.650 during the training stage and 0.853, 1.807, and 2.963, respectively during the testing stage. The performance of the Polynomial kernel function based GP model is better than the radial basis kernel function based GP model in terms of CC, RMSE, and MAE values are 0.853, 1.610, and 2.058, respectively for the training stage, and 0.839, 1.915 and 2.710, respectively for the testing stage. The agreement plot between the actual and predicted values of the VRUAFs using the GP_puk model for the testing stage is shown in Figure 9. All the predicted

Table 4 Results of Gaussian Processes using different Kernel Function

		Gaussian processes					
S. No	Kernel Function	TRAINING SET (146)			TESTING SET (72)		
		Correlation Coefficients	*Mean Absolute Error*	*Root Mean Square Error*	Correlation Coefficients	*Mean Absolute Error*	*Root Mean Square Error*
1	GP_RBFKernal	0.820	3.130	3.914	0.853	3.350	4.666
2	GP_Puk	0.885	1.231	1.650	0.853	1.807	2.963
3	GP_PolyKernel	0.853	1.610	2.058	0.839	1.915	2.710

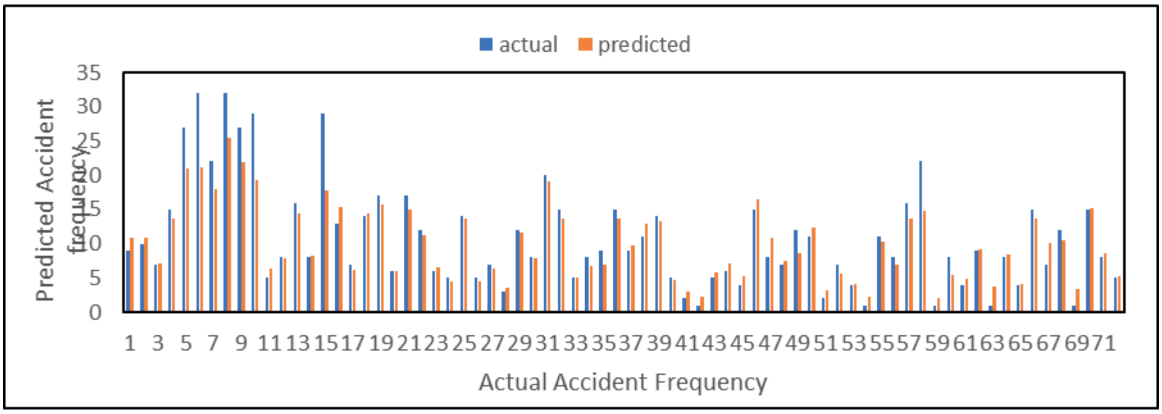

Figure 9 *GP Model comparison of actual and predicted values of VRUAF*

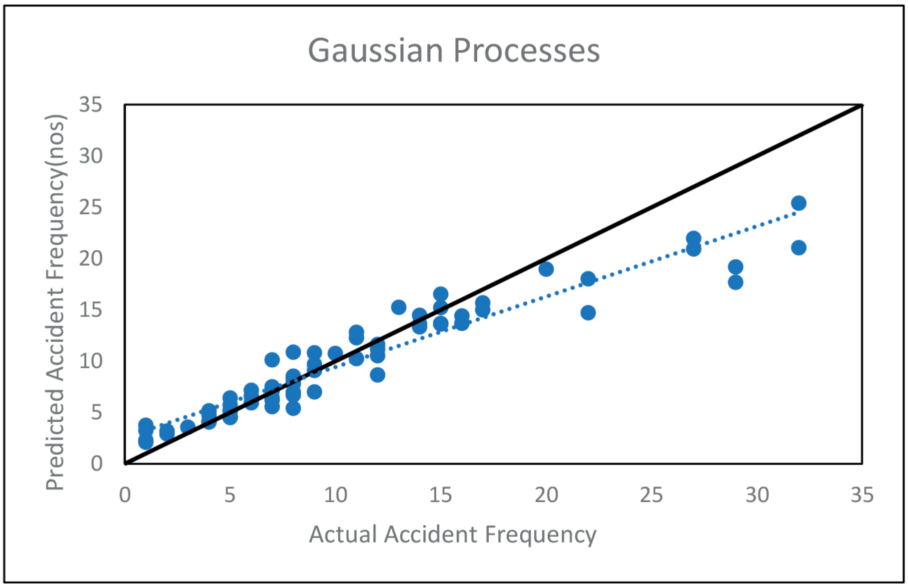

Figure 10 *Variation in scatter values with the actual values of the VRUAFs using the GP Model*

values lie closer to the line of perfect agreement (1:1) line. Figure 10 plots illustrate the distribution of actual VRUs accident frequency and predicted values of the VRUs accident frequency using the GP_puk model. The representation of this figure demonstrates that the GP_puk based model predicted the closer values to actual VRUs accident frequency values than the other models.

5 Comparison of selected artificial neural network model, Gaussian processes model, and support vector machine models

Table 5 summarizes the outcomes of the best selected ANN, SVM, and GP models. The results of this table indicate that the ANN-4 model was the most

Table 5 Performance evaluation parameters for different modelled VRUAF

Techniques	Training			Testing		
	CC	RMSE	MAE	CC	RMSE	MAE
ANN-4	0.924	1.124	1.458	0.912	1.198	1.799
SVM_puk	0.906	0.082	0.258	0.879	1.563	2.368
GP_Puk	0.885	1.231	1.650	0.853	1.807	2.963

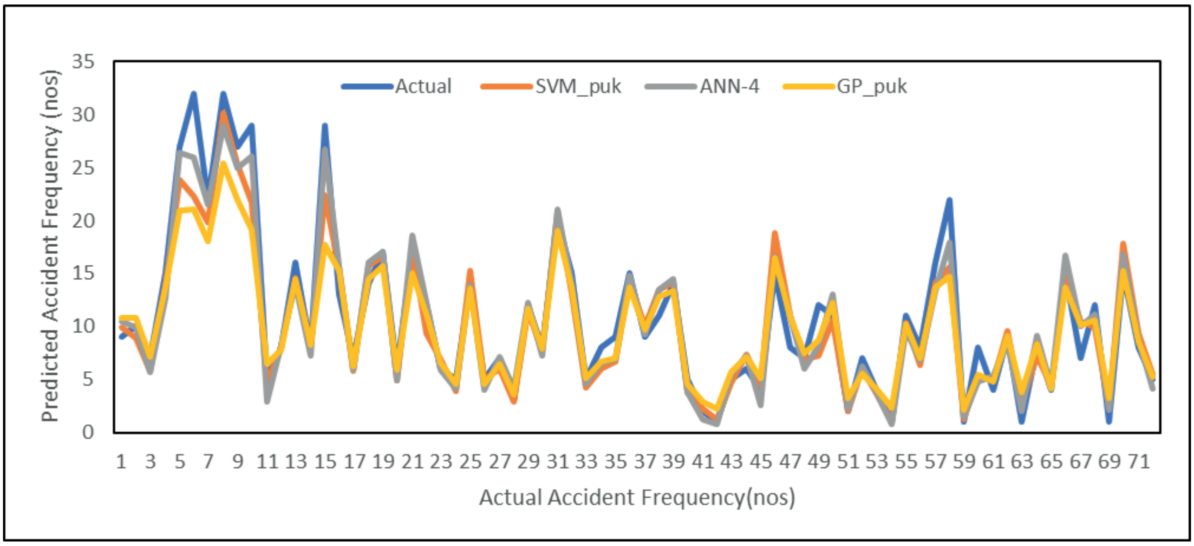

***Figure 11** Variation of the predicted values with the actual values of the VRUAFs using the SVM Model, GP Model, and ANN Model*

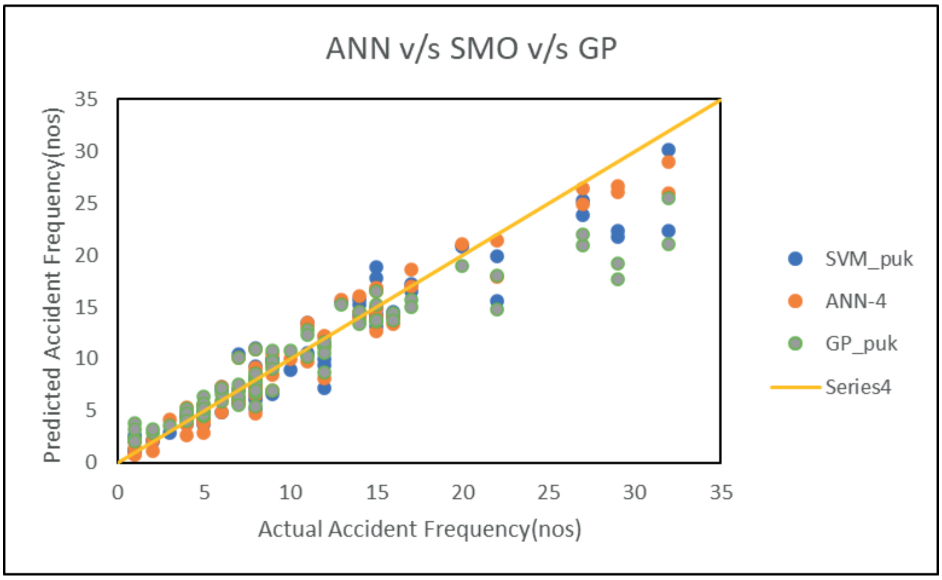

***Figure 12** Variation in scatter values with the actual values of the VRUAFs using SVM Model, GP Model, and ANN Model*

suitable model for accurately predicting the VRUAF with notably low CC and considerably high RMSE and MAE values.

The SVM Model, GP Model, and ANN Model comparison using actual and predicted VRUAF is shown in Figure 11. The ANN Model appears to be the best fit for predicting the VRUAF. The ANN-4 model lines for predicted values follow the same path as the actual values. The SVM, GP and ANN based best performing models are shown for actual and predicted VRUAF values in Figure 12. This figure illustrates that ANN-4 model points are much closer to the agreement line. Hence, the ANN-4 is the best performing model among all other applied models using this data set.

6 Conclusions

This investigation has assessed the performance of three regression-based modelling approaches to provide insights into the most suitable methods for predicting VRU (Vulnerable Road User) accident frequency. These approaches were ANN, SVM, and Gaussian Processes. This study employed various combination models as input variables, with VRUAF as the output. The ANN-4, SVM_Puk and GP_Puk model approach demonstrated exceptional efficiency when used with the ANN, SVM, and Gaussian Processes. According to the analysis, the ANN technique predicts the VRUAF more accurately than *the* GP and SVM. The R values for testing VRUAF data set come from the ANN approaches (0.904) are higher than GP (0.8539) and SVM (0.8795).

Similarly, the ANN model having values of RMSE and MAE (1.337 and 1.741) are much lower than the GP Model (1.807 and 2.963) and SVM model (1.563 and 2.368). Hence, ANN gave a more accurate VRUAF prediction than the GP and SVM models. Sensitivity investigation suggests that the most influential parameters in predicting vulnerable road user accident frequency, using the ANN model with the current data set, are LnL, LnADT, and Speed. Hence, the proper care should be taken when considering these parameters. The implementation of ANN is time-consuming, but it gives better accident prediction results than the actual happening of any accident. Hence, the proposed ANN model can be used to reduce the accident frequency and reduce accident costs and losses in terms of life and property.

Acknowledgment

The authors are thankful to the Public Works Department authorities, Rohtak, India, for the previous year's road geometry data, Toll Plaza National Highway-1, National Highway- 72, for road accident data. We are thankful to the Haryana police for recording road accidents on the online platform. This Harsamay online application collects the road accident data of various stretches.

Conflicts of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] BUBELINY, O., DADOVA, I., KUBINA, M., SOVIAR, J. The use of smart elements for the transport operation in the Slovak cities. *LOGI - Scientific Journal on Transport and Logistics* [online]. 2019, **10**, p. 51-60. eISSN 2336-3037. Available from: https://doi.org/10.2478/logi-2019-0015

[2] Road accidents in India [online]. 2021. Available from: www.morth.nic.in

[3] SIHAG, P., TIWARI, N. K., RANJAN, S. Prediction of cumulative infiltration of sandy soil using random forest approach. *Journal of Applied Water Engineering and Research* [online]. 2019, **7**(2), p. 118-142. eISSN 2324-9676. Available from: https://doi.org/10.1080/23249676.2018.1497557

[4] SINGH, B., SIHAG, P., TOMAR, A. SEHGAD, A. Estimation of compressive strength of high-strength concrete by random forest and M5P model tree approaches. *Journal of Materials and Engineering Structures JMES*. 2019, **6**(4), p. 583-892. eISSN 2170-127X.

[5] IVAZ, J., NIKOLIC, R. R., PETROVIC, D., DJOKOVIC, J. M., HADZIMA, B. Prediction of the work-related injuries based on neural networks. *System Safety: Human - Technical Facility - Environment* [online]. 2021, **3**(1), p. 19-37. eISSN 2657-5450. Available from: https://doi.org/10.2478/czoto-2021-0003

[6] PURI, D., KUMAR, R., SIHAG, P., THAKUR, M. S., PERVEEN, K., ALFAISAL, F. M., LEE, D. Analytical investigation of the impact of jet geometry on aeration effectiveness using soft computing techniques. *ACS Omega* [online]. 2023, **8**(42), p. 31811-31825. eISSN 2470-1343. Available from: https://doi.org/10.1021/acsomega.3c03294

[7] GOEL, A. K., KHAN, K., KUSHWAHA, A., SRIVASTAVA, V., MALIK, S., SINGH, A. A machine learning approach to analyze road accidents. In: 2022 IEEE International Conference on Blockchain and Distributed Systems Security ICBDS: proceedings [online]. IEEE. 2022. eISBN 978-1-6654-2832-3, p. 1-5. Available from: https://doi.org/10.1109/ICBDS53701.2022.9935867

[8] SUTHAR, M., AGGARWAL, P. Modeling CBR value using RF and M5P techniques. *Mendel* [online]. 2019, **25**(1), p. 73-78. ISSN 1803-3814, eISSN 2571-3701. Available from: https://doi.org/10.13164/mendel.2019.1.073

[9] SUN, Z., XING, Y., WANG, J., GU, X., LU, H., CHEN, Y. Exploring injury severity of vulnerable road user involved crashes across seasons: a hybrid method integrating random parameter logit model and Bayesian network. *Safety Science* [online]. 2022, **150**, 105682. ISSN 0925-7535, eISSN 1879-1042. Available from: https://doi.org/10.1016/j.ssci.2022.105682

[10] KOMOL, M. M. R., HASAN, M. M., ELHENAWY, M., YASMIN, S., MASOUD, M., RAKOTONIRAINY, A. Crash severity analysis of vulnerable road users using machine learning. *PLoS One* [online]. 2021, **16**, e0255828. eISSN 1932-6203. Available from: https://doi.org/10.1371/journal.pone.0255828

[11] BAMEL, K., DASS, S., JAGLAN, S., SUTHAR, M. Statistical analysis and development of accident prediction model of road safety conditions in Hisar city. *IOP Conference Series: Earth and Environmental Science* [online]. 2021, **889**, 012034. ISSN 1755-1315. Available from: https://doi.org/10.1088/1755-1315/889/1/012034

[12] LEE, J., MANNERING, F. Impact of roadside features on the frequency and severity of run-off-roadway accidents: an empirical analysis. *Accident Analysis and Prevention* [online]. 2002, **34**(2), p. 149-161. ISSN 0001-4575, eISSN 1879-2057. Available from: https://doi.org/10.1016/S0001-4575(01)00009-4

[13] MILTON, J., MANNERING, F. The relationship among highway geometrics, traffic-related elements and motor-vehicle accident frequencies. *Transportation* [online]. 1998, **25**, p. 395-413. ISSN 0049-4488, eISSN 1572-9435. Available from: https://doi.org/10.1023/A:1005095725001

[14] JAGLAN, S., AGGARWAL, P., KUMARI, S. A comparative study of M5P, ANN and RENB models for prediction of vulnerable road accident frequency, n.d.

[15] JAGLAN, S., KUMARI, S., AGGARWAL, P. Development of prediction models for vulnerable road user accident severity. *Optical Memory and Neural Networks* [online]. 2023, **32**, p. 346-363. ISSN 1060-992X, eISSN 1934-7898. Available from: https://doi.org/10.3103/S1060992X23040082

[16] JAGLAN, S., AGGARWAL, P., SINGHAL, D., DASS, S. Vulnerable road user accidents analysis on various roads of Haryana. *AIP Conference Proceedings* [online]. 2023, **2856**, 020003. ISSN 0094-243X, eISSN 1551-7616. Available from: https://doi.org/10.1063/5.0165909

[17] GARCIA DE SOTO, B., BUMBACHER, A., DEUBLEIN, M., ADEY, B. T. Predicting road traffic accidents using artificial neural network models. *Infrastructure Asset Management* [online]. 2018, **5**(4), p. 132-144. ISSN 2053-0242, eISSN 2053-0250. Available from: https://doi.org/10.1680/jinam.17.00028

[18] CODUR, M. Y., TORTUM, A. An artificial neural network model for highway accident prediction: a case study of Erzurum, Turkey. *Promet-Traffic and Transportation* [online]. 2015, **27**(3), p. 217-225. ISSN 0353-532, eISSN 1848-4069. Available from: https://doi.org/10.7307/ptt.v27i3.1551

[19] CORTES, C., VAPNIK, V. Support - vector networks. *Machine Learning* [online]. 1995, **20**, p. 273-297. ISSN 0885-6125, eISSN 1573-0565. Available from: https://doi.org/10.1007/BF00994018

[20] KARCH, J. D., BRANDMAIER, A. M., VOELKLE, M. C. Gaussian process panel modelling - machine learning inspired analysis of longitudinal panel data. *Frontiers in Psychology* [online]. 2020, **11**, 351. eISSN 1664-1078. Available from: https://doi.org/10.3389/fpsyg.2020.00351