International Journal of Computational and Experimental Science and ENgineering
(IJCESEN)

Vol. 11-No.2 (2025) pp. 3312-3321
http://www.ijcesen.com

Copyright © IJCESEN

ISSN: 2149-9144

Research Article

ANFIS-Based Traffic Accident Prediction Model for Karnataka State

Manoj P[1]*, Manjunath K C[2], Punith B. Kotagi[3]

[1] Department of Civil Engineering, The National Institute of Engineering, Mysore, affiliated to Visvesvaraya Technological University, Belagavi, Karnataka, India | Vidyavardhaka College of Engineering, Mysore, affiliated to Visvesvaraya Technological University, Belagavi, Karnataka, India
* **Corresponding Author Email:** manoj.p@vvce.ac.in - **ORCID:** 0000-0002-7185-5992

[2] Department of Civil Engineering, The National Institute of Engineering, Mysore, Karnataka, INDIA affiliated to Visvesvaraya Technological University, Belagavi, Karnataka, India
Email: kcmnie@gmail.com - **ORCID:** 0000-0002-0482-5910

[3] Department of Civil Engineering, The National Institute of Engineering, Mysore, affiliated to Visvesvaraya Technological University, Belagavi, Karnataka, India.
Email: punithbkotagi@nie.ac.in- **ORCID:** 0000-0001-5504-0534

Article Info:

DOI: 10.22399/ijcesen.2137
Received : 05 March 2025
Accepted : 07 May 2025

Keywords :

Accident prediction
Machine learning
ANFIS
Fuzzy subtractive clustering
Forecasting.

Abstract:

Given the seriousness of road traffic accidents as a public health concern, it is critical to comprehend the variables linked to an increase in the severity of injuries sustained by those who intervene in accidents. To improve road safety, make better decisions about road safety, and lessen the severity of crashes in the future, it is crucial to identify these elements. The study aimed to collect traffic data, analyse it to identify suitable variables for accident prediction, and to develop an accident predictive model suitable for Indian conditions. The dataset comprised of 67 blackspots, each containing 16 variables. One innovative aspect of this study involves the utilization of the fuzzy subtractive clustering algorithm to predict the accident. This approach holds theoretical promise in terms of computational efficiency. This contrasts with the traditional exponential increase in computational load with data dimensionality. Root Mean Squared Error (RMSE), and Coefficient of Determination (R^2) metrics were used to assess the model performance after the data was divided into training and validation sets, R2 of 0. 67 is obtained. The study emphasises the potential of machine learning to improve traffic safety.

1. Introduction

An estimated 1.19 million road traffic deaths occurred globally in 2021, or 15 deaths per 100,000 people, according to the World Health Organization's Global Status Report on Road Safety 2023. Road traffic injuries are the 12th largest cause of death for all age groups and continue to be the leading cause of death for those between the age of 5 and 29. Furthermore, there is a significant financial impact from traffic accidents; estimates put the worldwide cost as high as US$1.8 trillion, or 10–12% of the world's gross domestic product [1] . This poses a significant social and economic challenge. Road traffic accidents are the primary cause of death for children and young adults aged 5 to 29. Reducing the severity of collisions is a major global objective. Although all traffic accidents are concerning, vulnerable road users such as cyclists, motorcyclists, and pedestrians represent roughly 54% of all traffic fatalities globally and 43% in Europe [2]. According to Chang and Wang, the most significant factor influencing the severity of injuries is the category of vehicle, with susceptible drivers having the highest risk of suffering a serious injury or passing away [3]. Compared to car occupants, motorbike riders have a 30 times higher risk of dying in a traffic accident [4].

To increase the road safety, a thorough examination of traffic accidents and the identification of the variables influencing the severity of injuries sustained in the crash intervenient are essential. For statistical and machine learning techniques to effectively forecast the severity of injuries, the accessible data must be dependable and inclusive of the target population. After analysing data on traffic accidents in a few countries in southeast Europe, Laiou et al. [5] concluded that, to increase road safety in such nations, it is critical to improve the quality of the data sets. One of the most crucial elements in enhancing traffic safety is data, as without it, it would be impossible to gauge the effectiveness of the current road safety initiatives [6].

The comprehensive study involved identification and analysis of black spot in Karnataka State, using Adaptive neuro-fuzzy inference system.

2. Literature Review

Artificial neural networks are the most widely used machine learning technique for risk assessment in engineering, with Support Vector Machines (SVM) coming in second. Various techniques, including decision trees, Bayesian networks, and artificial neural networks, have been developed over the years to create models for predicting the severity of injuries in traffic accidents.

SVM and Ordered Probit (OP) models were used by Li et al. (2013) to analyse crash injury severity in China. This research was conducted on 1800 crash casualties. Several variables were taken into consideration, including the exit ramp's length and the shoulder width. 48.8% of the SVM model's predictions were accurate, compared to 44.0% for the OP model. Furthermore, it was found that the SVM model outperformed the OP model in terms of sensitivity and specificity [7]. Ibrahim and Far (2014) created a Real-Time Transportation Data Mining (R Trans Dmin) approach. This technique could be used to forecast future Real Time Accidents (RTA) related data and assess a set of real-time traffic data. The study focused on application of two decision tree types viz., Java 48 (J48) and Active Directory trees to 1385 accident records collected by the UK Department of Transport. The Active Directory tree approach registered 85.9%, while J48 registered 87.2%, based on the forecast accuracy statistics [8].

Mohamed (2014) forecasted the causes of traffic accidents based on 1000 real crashes in Dubai City using SVM with Gaussian Radial Basis function (RBF). The accuracy of this multi-class SVM model was over 75% [9]. To find the key factors affecting the severity prediction of crash datasets in Iran, Effati et al. (2015) used an integrated method to data mining that included SVM, coactive Neuro-Fuzzy, and ANN inference systems. This innovative integration technique registered good forecasts with an accuracy of 85.49% [10]. Perone (2015) used SVM, Logistic Regression, Random Forest, KNN, and Naïve Bayes to create a new prediction model to evaluate the severity of injuries in Brazil using 20798 accident records from the city of Porto Alegrers. According to African Union Commission (AUC) records, the highest satisfaction rates were achieved by SVM and Logistic Regression (94%), Random Forest (93%), KNN (90%), and Naïve Bayes (83%) [11].

To determine the key factors driving most RTAs, Sharma et al. (2016) collected 300 real accident cases from India and using Multi Layers Perceptron (MLP) and SVM (Gaussian kernel). Three categories were applied to their data set: 70% for training, 20% for cross-validation, and 10% for testing. LIBSVM, a support vector machine library, was integrated with Octave. The results of the investigation demonstrated that SVM with a Gaussian kernel function achieved 94% prediction accuracy, whereas conventional MLP achieved only 60% [12]. Gu et al. (2017) forecasted fatal traffic incidents in China using SVM. The goal of this study was to compare SVMs, KNNs, and Bayesian networks. The results showed that the particle swarm with mutation optimization-SVM-based traffic fatalities prediction model had a higher prediction precision (97%) and fewer mistakes (9%) in training and testing data. [13].

To determine the primary environmental features of RTA in the UK, Al-Radaideh and Daoud (2018) used ANN back propagation, SVM (polynomial Kernel), and Decision Trees (Random Forest C4.5/CART/J45). The experimental results of the study showed that decision trees, also known as random forests, generated the highest accurate forecast of accident severity in the United Kingdom (80.6%) [14].

Farhat et al. (2019) employed several data mining techniques, including decision trees and artificial neural networks (ANNs), to predict traffic accidents in Lebanon. The findings show that an artificial neural network (ANN) using a Multi Layers Perceptron (MLP) with two hidden layers and 42 neurones in each layer was the best strategy with an accuracy rate of prediction (94.6%) and an AUC (95.71%) [15].

Karthik et al. (2019) employed a range of data mining techniques, such as J48, Random Forest, and Naïve Bayesian, in the Thanjavur region of India to predict the primary causes of fatal accidents. Among other RTA parameters, ten years' worth of accident data were collected, including surface quality, road boundary, accident time, and accident location. At 56.96%, J48 had the highest accuracy rate, followed by Random Forest at 49% and Naïve Bayesian at 54% [16].

In addition to reviewing and summarising existing approaches to accident black spot identification, H. Cui et al. (2022) suggested a novel approach based on accident spacing distribution that circumvents road segmentation problems and yields more precise accident black spot identification [17].
To increase the road safety, a study by I. Karamanlis et. al(2023) identified road accident black spots in Northern Greece using both traditional and contemporary methods, such as binary logistic regression and different machine learning techniques [18].
The study conducted by A. Mbarek et al. (2023) suggests a paradigm for locating, categorising, and evaluating accident black spots on Moroccan rural roads by combining the Weighted Severity Index (WSI), Bagging Extreme Learning Machine (B-ELM), and Ordinal Regression (OR) approaches to examine how environmental elements and road conditions affect the severity of accidents [19].
The study conducted by Manoj et al. (2023) suggests that the Machine learning models like Random Forest and SVM are effectively used for predicting road accidents and identifying blackspots with high accuracy. The study stresses the need for real-time data integration using ML, IoT, and cloud technologies for proactive road safety.[20]

3. Objective and Methodology

3.1 Objective

The current study demonstrates how Adaptive neuro-fuzzy inference system (ANFIS) used to address the complex and vital issue of road safety. Overall, it demonstrates the multiple nature of research in forecasting and analysing the severity of traffic accidents. Through data-driven insights, the research contributes to the advancement of prediction model that may help avoid and mitigate the severity of traffic accidents.

3.2 Study Area and Data Collection

A road accident black spot is defined as a stretch of highway approximately 500 meters in length where, over the last three calendar years, either five road accidents involving fatalities or grievous injuries have occurred, or ten fatalities have taken place as per the guidelines of Ministry of Road Transport and Highways (MoRTH), 2019. The Planning and Road Asset Management Centre (PRAMC), a state government undertaking situated in Bangalore, Karnataka has identified several black spots across Karnataka State, India. However, 67 blackspots were considered for the study, and their locations are shown in Fig. 1.

Figure 1. *Locations of Selected Blackspots in State of Karnataka, India*

3.2.1 Primary Data*:* Primary data were collected through various surveys, including a road inventory survey and a spot speed study, with results recorded in an excel sheet for further analysis. The road inventory and detailed survey involved measuring road geometric details using a Total Station. Additional details within each black spot area were also documented. For traffic spot speed data, a radar speed gun was employed to measure vehicle speeds by detecting changes in frequency caused by the Doppler effect. Speed studies were conducted at all black spots, where a handheld radar gun was used specifically for two-wheelers and four-wheelers. The average spot speed was then calculated and utilized for modelling purposes.

3.2.2 Secondary Data: Secondary data for the study included First Information Reports (FIRs), vehicle insurance claim reports, and traffic volume data. FIRs, the primary source for accident data, were accessed with permission from the Superintendent of Police and gathered from police stations in Mysuru, Tumkuru, Ramanagara, Chamarajanagar, and Kodagu. These reports, covering 2017-2021, detailed incidents categorized by Indian Penal Code (IPC) sections 337, 338, and 304(A) for accidents involving various vehicles, including tempos, trucks, and buses. FIR copies from the designated black spot areas were also retrieved from the Karnataka State Police portal and contained specifics such as accident type, time, involved parties, causes, road condition, and vehicle type. IPC sections categorized accidents into minor (Section 337), major (Section 338), and fatal (Section 304(A)) categories.

Additionally, traffic volume data, including Passenger Car Unit (PCU) values, was obtained from the Office of the Chief Engineer (Communication and Buildings, South) in Bangalore, which provided information on traffic count for selected black spot stretches across districts.

3.3 Data Preparation

On two-lane undivided rural highways, black spots were examined with an emphasis on curved road segments only. Softwares viz., MS Excel, AutoCAD, E-Survey CAD were used to process survey data. Accidents were categorised as Minor, Major and Fatal using FIR data and Equivalent Accident Number (EAN) was calculated based on severity of accidents as per Indian Road Congress (IRC) 131-2022 under section 4.5.1.

A weighted total of accident severity, with severity weights obtained from IRC 131-2022 (minor = 2, major = 5, fatal = 10), was used to calculate EAN. More accurate accident analysis was made possible by this weighted system. The equation for calculating EAN is represented by below.

EAN = [(2 x minor accident) + (5 x major accident) + (10 x Fatal accident)]

The computed EAN for each Black Spot was then fixed as Dependent Variable (DV) and processed data consisting of 16 variables were considered as Independent Variables (IDV) for model development.

3.4 Selection of variables

3.4.1 List of variables

After processing the data, 17 variables were selected for model development and the statistical details of 16 Independent Variables (IDV) which is used for the prediction of EAN are presented in Table 1. For every variable, important statistics are given, including the mean, median, mode, standard deviation, minimum, and maximum. For instance, the speed at entry and exit varies very little, with a mean of about 65–66 m, and the mean radius is 126 m with a standard deviation of 96.9. To illustrate the distribution of these variables with their corresponding extremes, the dataset additionally contains measurements such as Length of Transition curve, Tangent Length, Carriage Way width and Shoulder width. Some variables, such as Passenger Car Unit (PCU) have a wide range of values (7109 to 25000), which indicate a wide range of observed conditions.

Table 1: *The statistical details of selected variables*

Sl. No.	Variables	Mean	Median	Mode	Standard deviation	Minimum	Maximum
1	**Radius [R] in m**	126	97.5	247	96.9	13.4	414
2	**Speed @ Entry [V(entry)] in Kmph**	65.3	66	66	3.07	58	72
3	**Speed @ Exit [V(exit)] in Kmph**	66.3	67	67	3.4	56	73
4	**Length of Transition Curve [Ls] in m**	46.9	26.2	10	60.5	6.05	382
5	**Tangent Length [L(tangent)] in m**	42.9	36.2	47.8	27.8	3.55	118
6	**Superelevation [e] in mm**	4.95	4.92	4.64	0.853	2.5	6.8
7	**Sight distance [SD] in m**	67.1	50	20	46.5	10	200
8	**Deflection Angle [D(angle] in m**	53.5	48	22.3	26	17.9	177
9	**Total width [TW] in m**	11.4	11	11	3.8	5.5	21
10	**Carriage Way Width [CW] in m**	8.22	7.5	5.5	2.69	3.75	15
11	**Shoulder width (Left) [SW(L) in m**	1.81	1.5	0	1.97	0	8.94
12	**Shoulder width (Right) [SW(R)] in m**	1.44	1.2	0	1.61	0	7
13	**Long Chord (LC) in m**	91.7	78.5	96.7	59.2	7.9	256
14	**Appex Distance (Es) in m**	12.5	8.16	4.77	12.6	1.07	77.3
15	**Mid Speed [V(mid)] in m**	50.5	52	56	10.3	20	67
16	**Passenger Car Unit [PCU]**	7109	5879	5462	3751	1769	25000
17	**Equivalent Accident Number [EAN]**	58.9	62	43	25.4	0	99

3.4.1 Correlation analysis

The correlation analysis of the variables used to estimate the Equivalent Accident Number (EAN) is shown in Fig. 2, emphasising important links and model predictors. The correlation matrix heatmap above displays the relationships between various variables related to road geometry, vehicle speed, sight distance, and accident data in terms of EAN (Equivalent Accident Number). The colour scale indicates the strength and direction of correlations, with dark red representing strong positive correlations and dark blue indicating strong negative correlations. For instance, "Radius" shows a strong positive correlation with "Sight distance" (0.88) and "Tangent Length" (0.83), suggesting that larger road radius may be associated with increased sight distance and tangent length. Conversely, "Sight distance" has a strong negative correlation with "EAN" (-0.79), implying that better visibility may reduce the frequency or severity of accidents. Similarly, "Speed at Entry" and "Speed at Exit" have moderate positive correlations with "EAN" 0.45 and 0.47, respectively, indicating that higher speeds at these points may contribute to more severe accidents. Designing safer road infrastructure may benefit from this analysis's and it facilitates in identifying the traffic and road features most frequently linked to accidents.

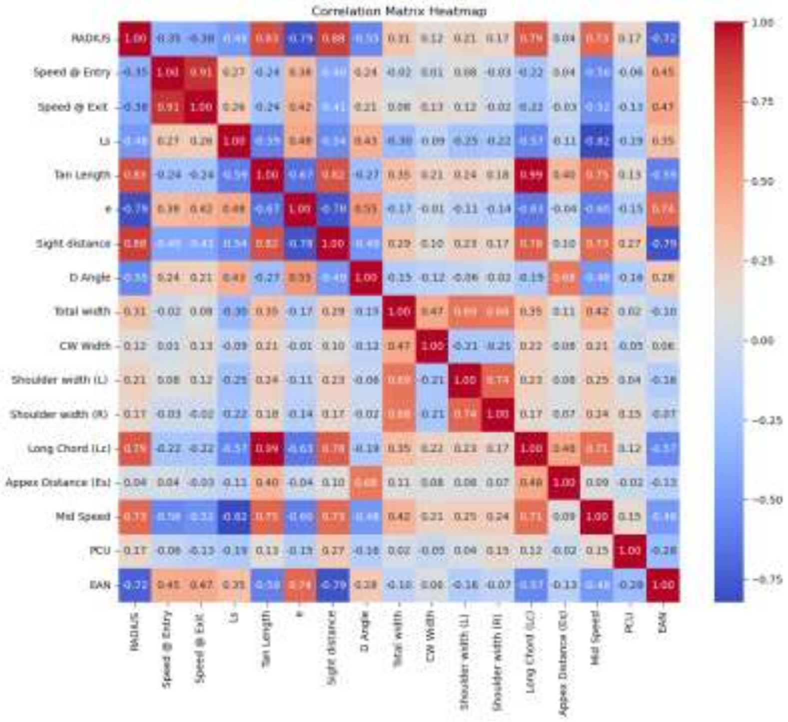

Figure 2. *Correlation between the variables*

Model Development

The dataset was divided into 80% testing and 20% validation for the application of ANFIS model.

Adaptive neuro-fuzzy inference system (ANFIS)

Fuzzy systems and artificial neural networks work together to generate a powerful hybrid system that can handle problems with complicated relationships. The constraints of fuzzy inference and ANN can be solved by ANFIS, one of the AI models. To develop a process that can handle complicated non-linear interactions between a collection of input and output, the ANFIS model combines the capabilities of ANN and fuzzy logic [20]. Figure 3 depicts the ANFIS's overall structure. First order Sugeno fuzzy has the following rule for a typical ANFIS, assuming the FIS has two inputs ('x' and 'y') and one output ('f').

Rule(1): $if\ \mu(x) is\ A_1\ and\ \mu(y) is\ B_1\ then\ f_1 = p_1 x + q_1 y + r_1$

Rule(2): $if\ \mu(x) is\ A_2\ and\ \mu(y) is\ B_2\ then\ f_2 = p_2 x + q_2 y + r_2$

A five-layer neural network configuration that followed the formulation and structure of ANFIS had membership function parameters for x and y inputs of A1, B1, A2, B1, and outlet function parameters of p1, q1, and r1, p2, q2, and r2[6].

ANFIS's architecture is composed of five levels. Numerous nodes represented by the node function are present in each layer. The parameter sets that are changeable in these nodes are represented by

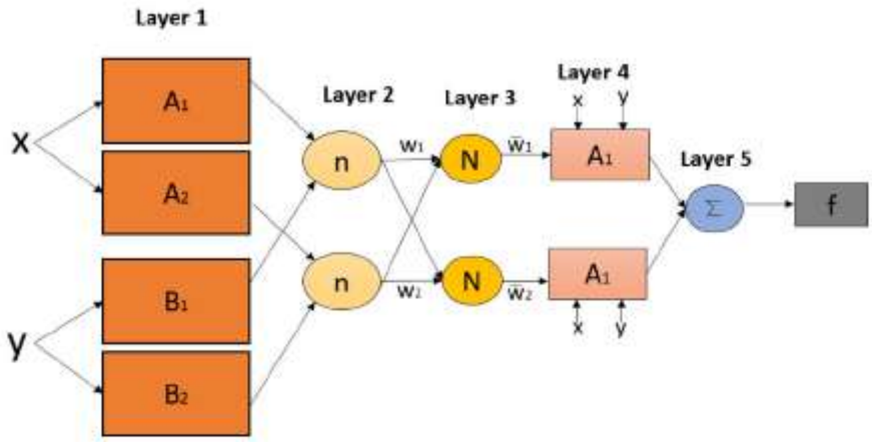

Figure 3. *ANFIS architecture*

adaptive nodes, which are indicated by squares,whereas the parameter sets that are fixed in the system are represented by fixed nodes, which are indicated by circles. The input for the current layer is made up of the output data from the nodes in the preceding layers. The figure below illustrates the general organisation of ANFIS. ANFIS's architecture is composed of five levels. Numerous nodes represented by the node function are present in each layer. The parameter sets that are changeable in these nodes are represented by adaptive nodes, which are indicated by squares, whereas the parameter sets that are fixed in the system are represented by fixed nodes, which are indicated by circles. The input for the current layer is made up of the output data from the nodes in the preceding layers.

To update the model parameters, ANFIS uses the hybrid-learning technique, which combines the "gradient descent" and "least squares" methods. This hybrid learning process consists of a forward pass and a backward pass for each epoch. The node output advances to layer 4 in the hybrid learning procedure's forward pass, where the ensuing parameters are determined using the least squares approach. The error signal propagates backwards during the backward pass, and gradient descent is used to update the premise parameters.

Total 16 input parameters of the ANFIS that are being taken into consideration in the current study. The EAN is the output. Data containing the chosen parameters for accident prediction were imported into the ANFIS using the Gaussian input parameter membership function.

Data were split into two groups for modelling: training data and testing data. The ANFIS was trained using the training data. The performance of the model was evaluated using the testing data. As a modelling tool for this work, the MATLAB Fuzzy Logic Toolbox ANFIS GUI was employed. Subtractive clustering (SC), a clustering algorithm, is used to create membership functions with a Gaussian shape automatically. This method creates if-then rules that are hazy. A quick, one-pass approach for determining the number of clusters and cluster centres in a piece of data is subtractive clustering. Iterative optimization-based clustering techniques and model identification techniques like ANFIS can both be initialised using the cluster estimates that were acquired. For the available data sets in this investigation, 3 cluster centres were identified. The number of fuzzy rule sets, each of which represents a different attribute of the cluster, would be equal to the number of cluster centres. The details on various parameters, and their values taken for modelling methods are given in Table 2.

Table 2: *Parameter values for clustering based ANFIS*

Parameters
Membership function = Gaussian Number of clusters = 3 Number of rules =45 Influence range= 0.5 Squash factor =1.25 Accepted ratio = 0.5 Rejected ratio = 0.15

Models' evaluation

We quantitatively assessed the models using five statistical indicators. Following is the calculation for these metrics.

Mean Square Error (MSE): The mean of the squared difference among the original and predicted values of the data set, as displayed in equation (1). It calculates the residuals' variance.

$$MSE = \frac{1}{N}\sum_{i=1}^{n}(y_i - \hat{y})^2 \dots\dots\dots(1)$$

Root Mean Square Error (RMSE): The mean square error's square root is RMSE equation (2). The standard deviation of the errors that happen when a prediction is made based on a dataset is known as RMSE.

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y})^2} \dots\dots\dots(2)$$

The coefficient of determination or R-squared: The percentage of the dependent variable's variation that the linear regression model can explain is indicated by the coefficient of determination, also known as R-squared. It is a scale-free score, therefore depending on the values little or huge, the R square value will be below one. It is calculated using the equation (3).

$$R^2 = 1 - \frac{\sum(y_i - \hat{y})^2}{\sum(y_i - \bar{y})^2} \dots\dots\dots\dots(3)$$

Where $\hat{y}$ is forecast value of y and $\bar{y}$ is average value of y.

4. Results and discussion

A hybrid method combining gradient descent back propagation and mean least squares optimisation methods is used by the ANFIS to work on the model and tune it. An error measure, which is the total of the squared difference between the output that occurs and what is desired, is decreased at each epoch. The overall EAN was calculated as a linear combination of the learnt values of the underlying premise parameters. Figure 4 depict the pattern of variance and distribution of the actual and predicted EAN for testing data, here, the blue dots represent real output while the red stars show EAN predictions.

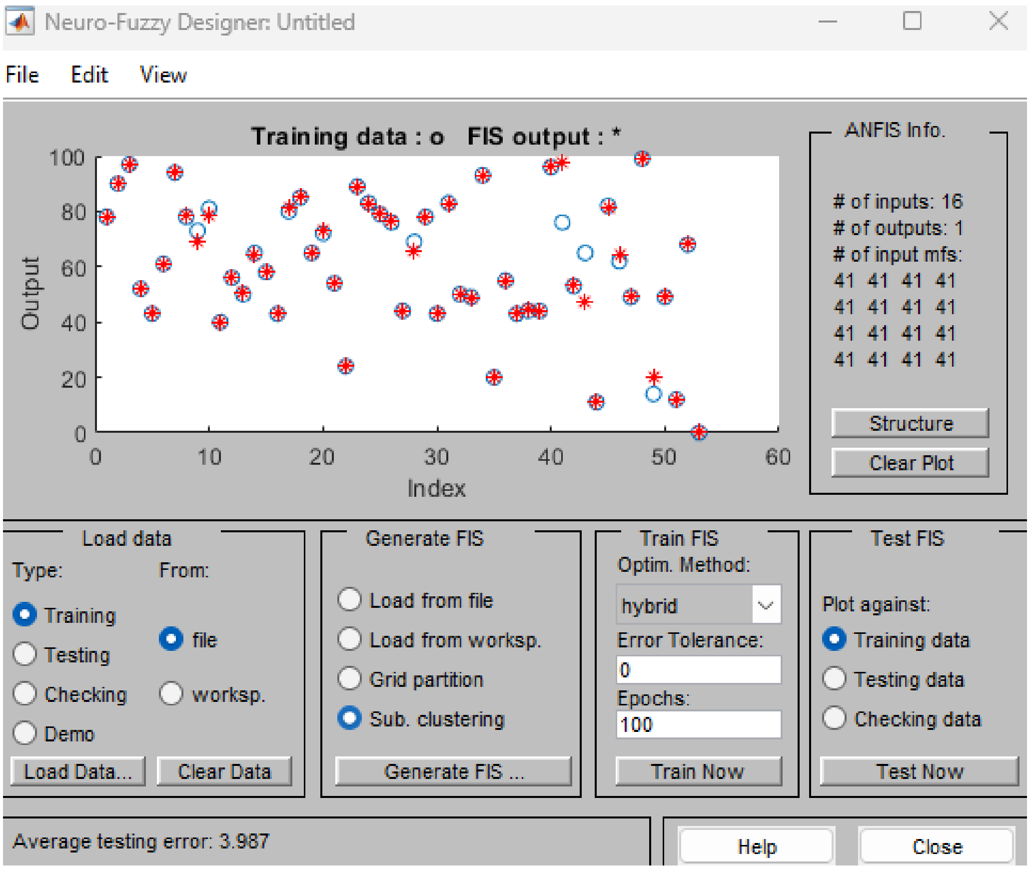

Figure 4. *Distribution of Actual and Predicted EAN*

The Figure 5 shows a Fuzzy Inference System (FIS) Rule Viewer from MATLAB, displaying how 16 input variables interact through fuzzy rules to produce an output. The yellow areas indicate active membership levels, helping visualize how each input contributes to the overall decision based on

41 rules. The scatter plots Figures 6 and 7 illustrate the performance of the accident prediction model during both training and validation phases. The training data yielded a lower R^2 value of 0.269, indicating a weaker correlation between predicted and actual values, possibly due to noise or overfitting limitations. In contrast, the validation data showed a significantly improved R^2 of 0.6828, reflecting a stronger fit and better generalization of the model to unseen data. The red dotted trendlines in both plots visually represent the relationship between predicted and actual outcomes, emphasizing the model's potential in real-world accident prediction scenarios.

Figure 5. *Set of Rules for the prediction of EAN*

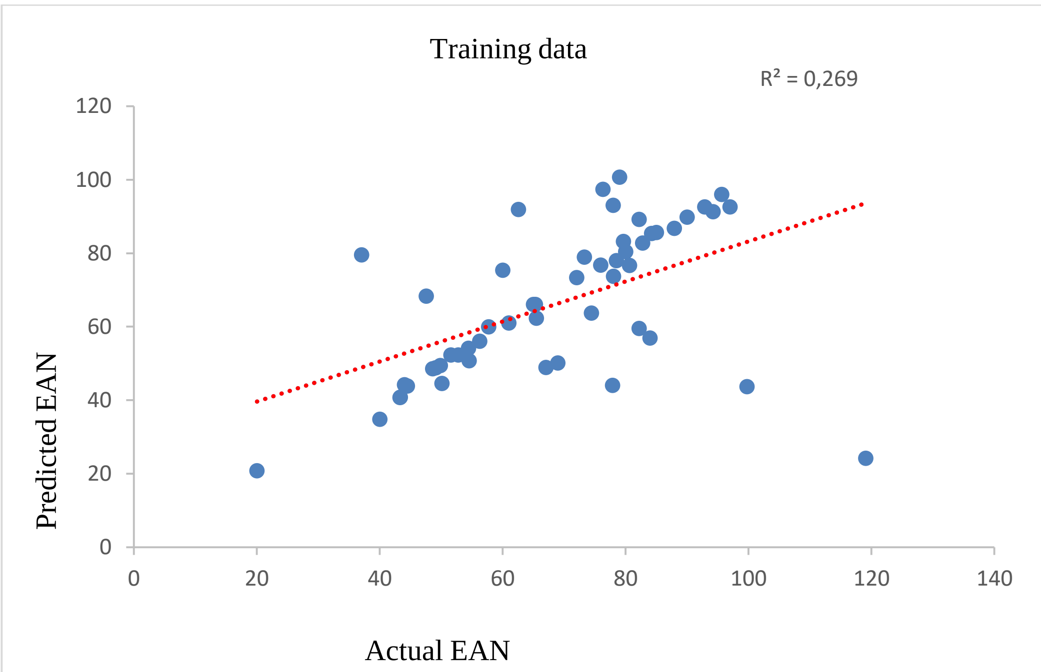

Figure 6. *Scatter plot of training data*

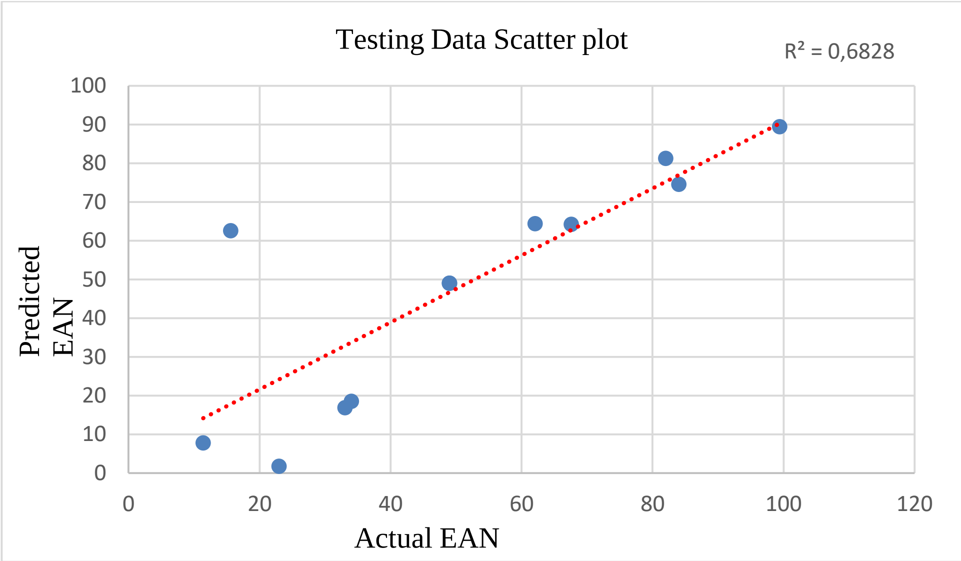

***Figure 7.** Scatter plot of testing data*

5. Conclusion

This study underscores the vital role of identifying key determinants influencing accident severity to strengthen road safety initiatives in Karnataka. Utilizing data from 67 blackspots, the application of fuzzy subtractive clustering and regression methods demonstrated the effectiveness of machine learning in predicting accident likelihood. With an R^2 value of 0.67 indicating strong predictive performance, the findings support the integration of intelligent data-driven models into traffic management frameworks. Such approaches can significantly enhance decision-making, enabling targeted safety interventions and contributing to evidence-based policy formulation.

Future research can expand this study by incorporating real-time traffic, weather, and vehicular data to enhance the predictive accuracy of the model. Integration of GIS and spatial analysis tools can help visualize accident-prone zones more effectively. Additionally, comparing the performance of other advanced machine learning models such as XGBoost, LSTM, or hybrid ensemble techniques may offer deeper insights. Collaboration with traffic enforcement agencies can enable the deployment of these models in intelligent transportation systems, leading to proactive accident prevention strategies. Lastly, extending the analysis to a wider range of geographic regions can improve the generalizability and robustness of the model.

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict 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
- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.
- **Author contributions:** The authors declare that they have equal right on this paper.
- **Funding information:** The authors declare that there is no funding to be acknowledged.
- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

[1] World Health Organization. (2023). *Global status report on road safety 2023: Summary*. World Health Organization.

[2] World Health Organization. (2019). *Global status report on road safety 2018*. World Health Organization.

[3] Wang, X., & Kim, S. H. (2019). Prediction and factor identification for crash severity: Comparison of discrete choice and tree-based models. *Transportation Research Record*, *2673*, 640–653.

[4] Lin, M.-R., & Kraus, J. F. (2008). Methodological issues in motorcycle injury epidemiology. *Accident Analysis & Prevention*, *40*, 1653–1660.

[5] Laiou, A., Papadimitriou, E., Yannis, G., & Milotti, A. (2017). Road safety data and information availability and priorities in South-East European regions. *Transportation Research Procedia*, *25*, 3703–3714.

[6] Sukor, N. A., & Sadullah, A. M. (2017). Addressing the road safety results impasse through an outcome-based approach in the state of Penang, Malaysia. *Engineering Heritage Journal*, *1*, 21–24.

[7] Li, Z., Liu, P., Wang, W., & Xu, C. (2012). Using support vector machine models for crash injury severity analysis. *Accident Analysis & Prevention*, *45*, 478–486.

[8] Ibrahim, H., & Far, B. H. (2014). Data-oriented intelligent transportation systems. In *Proceedings of the 2014 IEEE 15th International Conference on Information Reuse and Integration (IEEE IRI 2014)* (pp. 322–329).

[9] Mohamed, E. A. (2014). Predicting causes of traffic road accidents using multi-class support vector machines. *Journal of Communication and Computer*, *11*, 441–447.

[10] Effati, M., Thill, J.-C., & Shabani, S. (2015). Geospatial and machine learning techniques for wicked social science problems: Analysis of crash severity on a regional highway corridor. *Journal of Geographical Systems*, *17*, 107–135.

[11] Perone, C. S. (2015). Injury risk prediction for traffic accidents in Porto Alegre/RS, Brazil. *arXiv preprint*, arXiv:1502.00245.

[12] Sharma, B., Katiyar, V. K., & Kumar, K. (2016). Traffic accident prediction model using support vector machines with Gaussian kernel. In *Proceedings of Fifth International Conference on Soft Computing for Problem Solving: SocProS 2015, Volume 2* (pp. 1–10).

[13] Gu, X., Li, T., Wang, Y., Zhang, L., Wang, Y., & Yao, J. (2018). Traffic fatalities prediction using support vector machine with hybrid particle swarm optimization. *Journal of Algorithms & Computational Technology, 12*, 20–29.

[14] Al-Radaideh, Q. A., & Daoud, E. J. (2018). Data mining methods for traffic accident severity prediction. *International Journal of Neural Networks and Advanced Applications, 5*, 1–12.

[15] Farhat, Z., Karouni, A., Daya, B., & Chauvet, P. (2019). Comparative study between decision trees and neural networks to predict fatal road accidents in Lebanon. In *5th International Conference on Computer Science, Information Technology* (pp. 1–14).

[16] Karthik, D., & Vijayarekha, K. (2020). Identifying efficient road safety prediction model using data mining classifiers: Recent advancements.

[17] Cui, H., Dong, J., Zhu, M., Li, X., & Wang, Q. (2022). Identifying accident black spots based on the accident spacing distribution. *Journal of Traffic and Transportation Engineering (English Edition), 9*, 1017–1026.

[18] Karamanlis, I., Kokkalis, A., Profillidis, V., Botzoris, G., & Galanis, A. (2023). Identifying road accident black spots using classical and modern approaches. *WSEAS Transactions on Systems, 22*, 556–565.

[19] Mbarek, A., Jiber, M., Yahyaouy, A., & Sabri, A. (2023). Black spots identification on rural roads based on extreme learning machine. *International Journal of Electrical and Computer Engineering, 13*, 3149–3160.

[20] Manoj, P. (2023). Enhancing road safety through blackspots mitigative measures – A review. *Sigma Journal of Engineering and Natural Sciences – Sigma Mühendislik ve Fen Bilimleri Dergisi*. https://doi.org/10.14744/sigma.2023.00077