Comparative Study of Machine Learning Algorithm for Traffic Accident Prediction and Prevention

Manisha Verma and Bharti Sharma

August 30, 2024

Comparative Study of Machine Learning Algorithm for Traffic Accident Prediction and Prevention

Manisha Verma[1], Dr. Bharti Sharma[2]

Research Scholar[1], Associate Professor[2]

Department of Computer Application, DIT University, Dehradun, Uttarakhand, India

Abstract

The goal of this research is to compare supervised learning algorithms in rural and urban areas. Road accidents have become a major concern and the main cause of mortality globally. The objective of this research article is to analyse road accidents, estimate the severity of each accident using supervised learning algorithms, and identify the elements that contributed to these accidents. The goal is to solve the issue of safety by creating an accurate prediction model capable of identifying trends in diverse settings and preventing traffic accidents. To achieve this, machine learning algorithms are employed to predict different traffic accident scenarios and identify the most significant factors contributing to these accidents. By utilizing a machine learning model, a cost-effective approach can be developed for implementing safety measures. The ultimate goal of this model is to enhance the accuracy of accident prevention measures and improve overall security. For the analysis, three supervised learning algorithms, namely random forest, decision tree, and SVM, are utilized to predict traffic accident data. These algorithms are chosen due to their ability to provide accurate predictions while considering low-budget scientific measures for accident reduction. In conclusion, this research paper aims is to compare the performance of different supervised machine learning algorithms in rural and urban environments in order to analyze road accidents, determine accident severity, and identify the factors responsible for these accidents. By utilizing machine learning-based models and predictive algorithms, the goal is to develop cost-effective safety measures for preventing traffic accidents and improving overall security. Through comprehensive analysis and evaluation, this research aims to contribute to the field of accident prevention and safety enhancement.

I. Introduction

A considerable negative effect of transportation systems are harms, victims, economic losses and traffic congestion due to road accidents in rural and urban areas. The traffic accidents (TAs) have become a major problem due to increase the number of transportation on the road in cities and villages [8][9]. This research work is based on traffic accident prediction using ML algorithms as a category of supervised learning algorithms in rural and urban areas for safety to transportation. The challenging tasks of prediction to large amount of traffic flow data due to the involved in accidents [13][14]. The purpose of prediction of traffic accidents many machine learning algorithms are used for accuracy. The machine learning algorithm is used for solve the issue of accuracy for prediction [15]. The aim of this research paper is improve the traffic accident predictions accuracy in large and complex datasets using ML algorithms. In this research, use three algorithms as decision tree, support vector machine (SVM) and random forest of supervised learning which category of machine learning for traffic data prediction in road network. The dataset in a training model also contains parameters of labeled input and output under supervised learning model. In supervised learning based model use for both purpose as validation and training of model [16][17]. The supervised learning model is classified into two type model as classification model and regression model [18]. In the classification model, the output variable is categorical with classify the purpose of new data into pre-defined categories of data [19][20]. In regression model, output variable is based on input variable which predict continuously in model [21][22]. The three main algorithms as SVM, random forest and decision tree

which category of classification and regression model are used in this research work. The support vector machine (SVM) is use for find the hyper plan in the space of N-dimension which classifies in the data points [23]. The numbers of features in the dataset are the number of dimensions in hyper plane in SVM [24]. The Decision tree is a hierarchical tree structure model in which consist internal nodes presents features, rule presented by branches and algorithm results presented by leaf nodes [25]. The random forest algorithm is applied for solve to complex problems using multiple decision tree which produce single output [3][4]. The structure of traffic accident prediction is show in figure 1 as:
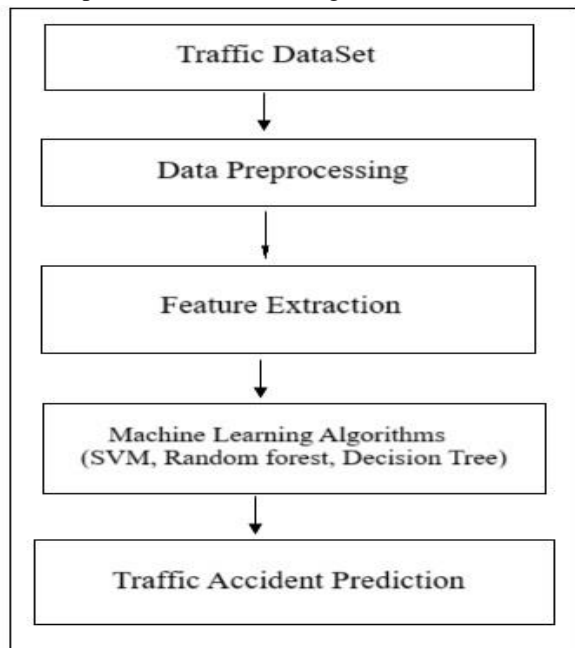

Figure 1.Prediction Model : Traffic Accident

II.   Literature review

Megnidio-Tchoukouegno M. & Adedayo Adedeji J., (2023), Proposed the research on ML based road traffic accident prediction using three data driven methodologies as prediction of future traffic accidents, identifier of traffic variables and technical inspection of vehicle in transportation society. The outcome of this research is the performance of decision tree (DT) algorithm is higher than XGBoost and LightGBM under traffic accident prediction [1].

Ardakani, S. P. et al., (2023), Proposed the predictive model for traffic accidents reduce using different ML techniques as multinomial logistic regression decision tree, naïve Bayesian classification and random forest. The result of this work is 80% accurate for prediction [2].

Azhar A., et al., (2023) , Proposed the prediction model for detect and predict the traffic accidents using deep learning techniques on behalf of social media posts. The result of this work is 94.2% traffic accident detection and 75% accuracy for prediction using deep learning models [3].

Azhar A., et al., (2023), Proposed the prediction model for detect and predict the traffic accidents using deep learning techniques on behalf of social media posts. The result of this work is 94.2% traffic accident detection and 75% accuracy for prediction using deep learning models [3].

Akhila N. et al., (2023), Proposed the review on traffic flow prediction with performance comparison between supervised and deep learning approaches. The outcome of this paper is deep learning is superior then supervised learning approach [4].

Shuang, Q. & Zhang Z. (2023), Proposed the analysis of relationship between different fatal accidents causes prediction using optimize machine learning model in construction industry. The outcome of this work is F1 score of proposed model is higher than other model [5].

Du L. et al., (2023), Proposed the study of methodology for predict and classify to illegal dumping in case study of south Australia. The methodology is combining to geographical information system and machine learning for risk areas in traffic conditions. The outcome of this study is the proposed theoretical model is practical implemented on large areas of illegal dumping management [6].

Prajapati G. et. al., (2023), Proposed the prediction model for traffic accident prediction using Kaggle's dataset of traffic accidents with different technique of machine learning as SVM, decision tree and random forest. The result of this work is 98% accuracy of prediction of traffic accidents in pinpoints hotspots [7].

Berhanu Y. et al., (2023), Proposed the approaches for prevents to accidents in congestion of traffic using different prediction techniques of machine learning as

decision tree, logical regression, random forest, KNN, DNN and SVM with GIS based spatial methods and advanced optimization algorithm. The outcome of this work is proposed model is accurate for congestion reduce in traffic and accidents identification [9].

Gangwani D. & Gangwani P., (2023), Proposed the application of ML and AI in intelligent transportation system (ITS) in road network for handle the issues of road accidents, traffic congestion and road anomalies. The outcome of this study is ML algorithms is used for predict to traffic accidents for safety reason [10].

Ammar S. et al., (2023), Proposed the review in data driven approach with predict traffic conditions, monitoring, for road safety and focus for analysis of accident detection with prediction and accident prevention. The outcome of this survey is data driven approach is used in multidisciplinary work in road network for traffic accident detection and prevention [11].

Cheng L. et al., (2023), Proposed the approach of data driven using social media data like post, replies and likes for predict the stock trading behavior using various machine learning methods as decision tree, random forest, SVM and KNN. The outcome of this work is rumor propagation is outer performers than other variables of abnormal trading behavior prediction [12].

III. Proposed work

The prediction of actual traffic data is a common problem due to traffic congestion. The issue of real time traffic data prediction is difficult on behalf of past and recent dataset but identify the pattern of traffic congestion with the help these datasets. In urban and rural areas, traditional method is not sufficient for handle complex traffic patterns. The purpose of this prediction is for gain the real time information in traffic congestion. The propose work of this research paper is observe the performance of three supervised learning algorithms as SVM, decision tree and random forest in traffic accident prediction under the congestion of traffic using three different performance parameters as accuracy, precision and recall.

IV. System Design

The design of system for this research is based on variables of traffic prediction which is date, time and junction in traffic congestion. The data is collected from traffic congestion and built prediction model for traffic accident prediction using different library of python as pandas, numpy, keras and sklearn library with using classification and regression model. The traffic dataset is collected from number of vehicles. The algorithms of ML are used for prediction to traffic accidents in road network. The figure 2 shows the outline of the traffic accident prediction in this paper as:
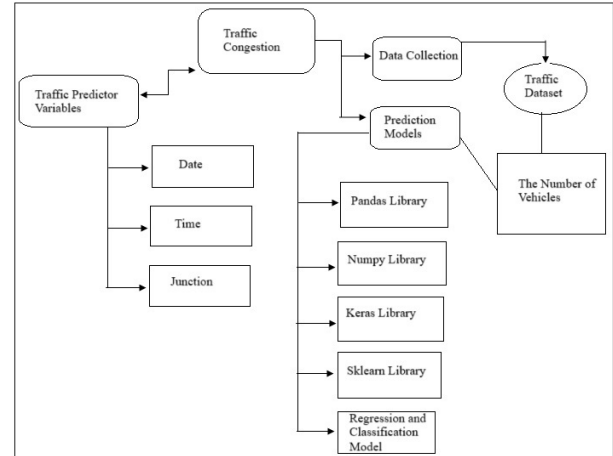

Figure 2.Traffic Accident prediction system

V. Research Methodology

The research methodology of this research paper as:

a) Data set: The dataset is collected from Kaggle website. The information of dataset contains day, date, coded day, zone, weather, temperature and traffic.

b) Data preprocessing: The predictive analysis process is use for remove or deletes the irrelevant data using machine learning algorithm.

c) Modeling strategies: In modeling strategy, classification and regression model is use for prediction to number of vehicles, number of causalities and number of accident strictness.

d) Proposed algorithms: In this research work, three algorithm of supervised learning as SVM, decision tree and random forest are used for prediction to traffic accidents.

e) Performance matrices: There are three performance parameters as precision, recall and accuracy is used for evaluation of performance under traffic accident prediction.

Precision: This is the ratio of true positive and total of true positive and false positive.

$$Precision = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

Recall: This is the ratio of true positive and total of true positive and false negative.

$$Recall = \frac{TP}{TP + FN}$$

Accuracy: Is the ratio of total of true positive and true negative and sum of true and false positive and negative values.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

VI. Analysis and Results

In this section, implement of classification and regression model based algorithms as random forest, SVM and decision tree for traffic accident prediction using traffic dataset under different performance parameters as accuracy, precision and recall.

1. Dataset: The dataset of traffic accident prediction model is show in table 2 as:

Day	Date	CodedDay	Zone	Weather	Temperatu	Traffic
Friday	18/11/202	4	2	34	18	2
Friday	18/11/202	4	3	37	17	3
Friday	18/11/202	4	4	29	24	6
Friday	18/11/202	4	5	22	22	9
Friday	18/11/202	4	6	18	45	7
Friday	18/11/202	4	7	10	18	2
Friday	18/11/202	4	8	45	27	3
Friday	18/11/202	4	9	38	16	3

Table 2.Dataset of prediction

2. Analysis and Result: In this section, discuss the results after analysis of performance of algorithms for prediction to traffic accidents in urban and rural areas. The table 3 shows the results of performance after analysis of SVM, decision tree and random forest algorithm for traffic accident prediction.

Algorithms	Accuracy	Precision	Recall
SVM	88%	86.87%	80%
Decision Tree	89%	87.86%	81%
Random Forest	92%	89.81%	83%

Table 3.Performance results

When analysis of proposed algorithms for prediction under three different performance parameters as accuracy, precision and recall the outcome is random forest algorithm is superior then another two algorithms as SVM and decision tree.

Following figure 3 show the graph of accuracy for traffic accident prediction using algorithms. In accuracy, the algorithm Random Forest is accurate than decision tree and SVM algorithm means random forest is superior for traffic accident prediction with compare to SVM and decision tree. The DT algorithm is best for prediction with compare to SVM but less accurate to random forest algorithm.
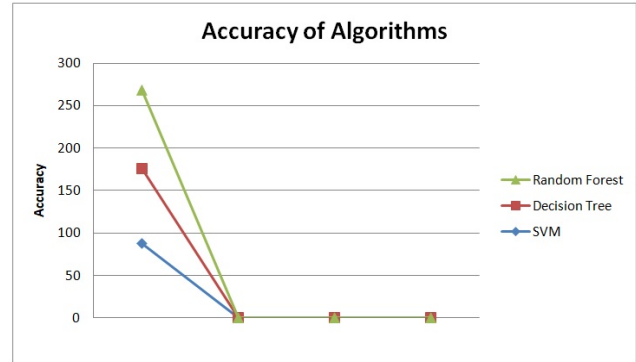

Figure 3.Prediction of accuracy

Figure 4 show the precision graph of algorithm. If analysis of performance under precision parameter, the random forest algorithm is greater than SVM and decision tree due to reduce congestion and easily prediction to traffic accidents. The SVM algorithm is not superior to decision tree and random forest algorithm under precision.
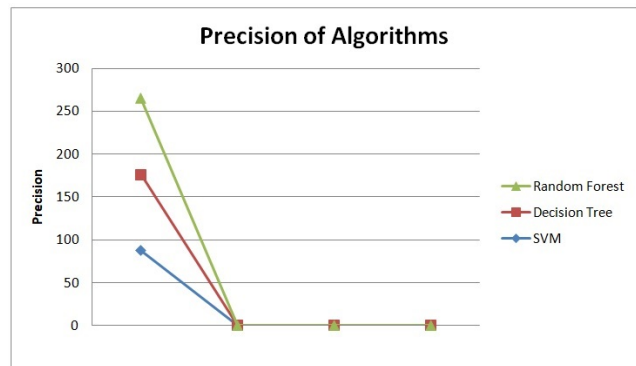

Figure 4.Prediction of precision

Figure 5 show the prediction graph of proposed algorithms under recall. In analysis of recall, the decision tree algorithm is best than SVM but less to random forest algorithm. The random forest algorithm is higher than decision tree and SVM algorithm in analysis for prediction under recall.
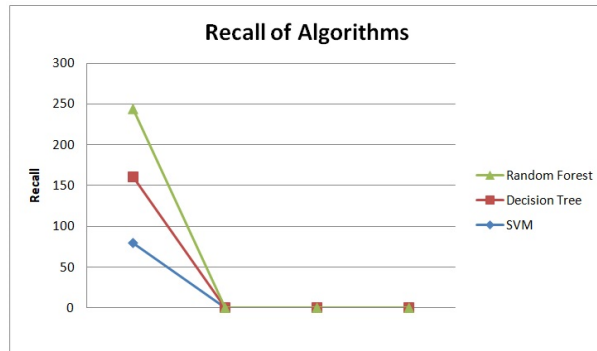
Figure 5.Prediction to recall

VII. Conclusion and Future scope

In this paper, use a machine learning algorithm for predicts to traffic accidents by using a model of classification and regression. This system is use for traffic accident control using traffic condition. This system presents the information of the number of vehicle passing through a particular intersection to road condition. The data of traffic accident prediction is recognizing by traffic condition. This paper present the comparison of three supervised learning algorithm as SVM, decision tree and random forest under the performance of traffic accident prediction using three performance matrices as accuracy, precision and recall. In this research work, use classification and regression model for built propose predictive model. The outcome of this research is random forest algorithm is superior to other two algorithms as SVM and decision tree. The aim of this research paper is predict the data of traffic accidents in traffic congestion. The future scope of this research, the improvement of traffic prediction using ANN, deep learning and big data due to improve the accuracy of prediction with proper suggestion of route of their destination. The further work of this research is built new forecasting model for improve the accuracy of prediction in transportation system.

References

[1]   Mireille Megnidio-Tchoukouegno and Jacob Adedayo Adedeji. "Machine learning for road traffic accident improvement and environmental resource management in the transportation sector." Sustainability 15, no. 3 (2023): 2014.

[2]   Saeid Pourroostaei Ardakani, Xiangning Liang, Kal Tenna Mengistu, Richard Sugianto So, Xuhui Wei, Baojie He, and Ali Cheshmehzangi. "Road Car Accident Prediction Using a Machine-Learning-Enabled Data Analysis." Sustainability 15, no. 7 (2023): 5939.

[3]   Anique Azhar, Saddaf Rubab, Malik M. Khan, Yawar Abbas Bangash, Mohammad Dahman Alshehri, Fizza Illahi, and Ali Kashif Bashir. "Detection and prediction of traffic accidents using deep learning techniques." Cluster Computing 26, no. 1 (2023): 477-493.

[4]   N. Akhila, M. Kavya, M. Soumith Reddy, and Prasanta Kumar Pradhan. "Traffic Prediction For Intelligent Transport System Using Machine Learning." (2023).

[5]   Qing Shuang, and Zerong Zhang. "Determining Critical Cause Combination of Fatality Accidents on Construction Sites with Machine Learning Techniques." Buildings 13, no. 2 (2023): 345.

[6]   Linwei Du, Jian Zuo, John Vanzo, Ruidong Chang, and George Zillante. "Assessing and predicting the illegal dumping risks in relation to road characteristics." Waste Management 169 (2023): 332-341.

[7]   Gaurav Prajapati, Lav Kumar, and Smt Rekha S. Patil. "Road Accident Prediction Using Machine Learning." Journal of Scientific Research and Technology (2023): 48-59.

[8]   Dhruv Vimal Shah. "Prediction of Accident Severity Using Machine Learning Algorithms." PhD diss., Dublin, National College of Ireland, 2023.

[9]   Yetay Berhanu, Esayas Alemayehu, and Dietrich Schröder. "Examining Car Accident Prediction Techniques and Road Traffic Congestion: A Comparative Analysis of Road Safety and Prevention of World Challenges in Low-Income and High-Income Countries." Journal of Advanced Transportation 2023 (2023).

[10]   Divya Gangwani and Pranav Gangwani. "Applications of machine learning and artificial intelligence in intelligent transportation system: A review." Applications of Artificial Intelligence and Machine Learning: Select Proceedings of ICAAAIML 2020 (2021): 203-216.

[11]   Sohail Ammar, Muhammad Aamir Cheema, Mohammed Eunus Ali, Adel N. Toosi, and Hesham A. Rakha. "Data-driven approaches for road safety: a comprehensive systematic literature review." Safety science 158 (2023): 105949.

[12]   Li-Chen Cheng, Wei-Ting Lu, and Benjamin Yeo. "Predicting abnormal trading behavior from internet

rumor propagation: a machine learning approach." Financial Innovation 9, no. 1 (2023): 3.

[13] Mauro Caetano. "Aviation accident and incident forecasting combining occurrence investigation and meteorological data using machine learning." Aviation 27, no. 1 (2023): 47-56.

[14] Ningzhe Xu, Ruggiero Lovreglio, Erica D. Kuligowski, Thomas J. Cova, Daniel Nilsson, and Xilei Zhao. "Predicting and Assessing Wildfire Evacuation Decision-Making Using Machine Learning: Findings from the 2019 Kincade Fire." Fire Technology 59, no. 2 (2023): 793-825.

[15] Vidya GS and Hari VS. "Prediction of bus passenger traffic using Gaussian process regression." Journal of Signal Processing Systems 95, no. 2-3 (2023): 281-292.

[16] Dhyani Dhaval Desai, Joyeeta Dey, Sandeep Kumar Satapathy, Shruti Mishra, Sachi Nandan Mohanty, Pallavi Mishra, and Sandeep Kumar Panda. "Optimal Ambulance Positioning for Road Accidents with Deep Embedded Clustering." IEEE Access (2023).

[17] Seyed Mohammad Hossein Moosavi, Mahdi Aghaabbasi, Choon Wah Yuen, and Danial Jahed Armaghani. "Evaluation of Applicability and Accuracy of Bus Travel Time Prediction in High and Low Frequency Bus Routes Using Tree-Based ML Techniques." Journal of Soft Computing in Civil Engineering 7, no. 2 (2023): 74-97.

[18] Arti Deshpande and Arya Kumar. "Workplace Incident and Injuries Prevention Using Machine Learning." In International Conference on Applications of Machine Intelligence and Data Analytics (ICAMIDA 2022), pp. 499-512. Atlantis Press, 2023.

[19] Md Kamrul Islam, Imran Reza, Uneb Gazder, Rocksana Akter, Md Arifuzzaman, and Muhammad Muhitur Rahman. "Predicting road crash severity using classifier models and crash hotspots." Applied Sciences 12, no. 22 (2022): 11354.

[20] Zhaoshi Geng, Xiaofeng Ji, Rui Cao, Mengyuan Lu, and Wenwen Qin. "A conflict measures-based extreme value theory approach to predicting truck collisions and identifying high-risk scenes on two-lane rural highways." Sustainability 14, no. 18 (2022): 11212.

[21] Samuel Olugbade, Stephen Ojo, Agbotiname Lucky Imoize, Joseph Isabona, and Mathew O. Alaba. "A review of artificial intelligence and machine learning for incident detectors in road transport systems." Mathematical and Computational Applications 27, no. 5 (2022): 77.

[22] Giuliano Ramírez, Kevin Salazar, Vicente Barria, Oscar Pinto, Lilian San Martin, Raúl Carrasco, Diego Fuentealba, and Gustavo Gatica. "Accident risk detection in urban trees using machine learning and fuzzy logic." Procedia Computer Science 203 (2022): 471-475.

[23] Chukwuebuka Joseph Ejiyi, Zhen Qin, Abdulhaq Adetunji Salako, Monday Nkanta Happy, Grace Ugochi Nneji, Chiagoziem Chima Ukwuoma, Ijeoma Amuche Chikwendu, and Ji Gen. "Comparative analysis of building insurance prediction using some machine learning algorithms." (2022).

[24] Jing Wang, and Filip Biljecki. "Unsupervised machine learning in urban studies: A systematic review of applications." Cities 129 (2022): 103925.

[25] Omolola Faith Ademola, Sanjay Misra, and Akshat Agrawal. "Improving Real-Time Intelligent Transportation Systems in Predicting Road Accident." In The International Conference on Recent Innovations in Computing, pp. 225-239. Singapore: Springer Nature Singapore, 2022.