International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307
OPEN ACCESS
Available Online at : www.ijsrcseit.com
doi : https://doi.org/10.32628/CSEIT241024

Road Accident Severity Detection In Smart Cities

Deeksha K, Kavya S, Nikita J, Evangeline R. C
Information Science, Nitte Meenakshi Institute of Technology, Bengaluru, Karnataka, India

ARTICLEINFO

Article History:

Accepted: 05 March 2024
Published: 16 March 2024

Publication Issue
Volume 10, Issue 2
March-April-2024

Page Number
180-187

ABSTRACT

Ensuring safety, in cities is a focus in the development of urban areas requiring new and creative methods for categorizing and managing accidents. Traditional approaches often face challenges in evaluating accident seriousness within changing city environments. This research utilizes Long Short Term Memory (LSTM) and Convolutional Neural Network (CNN) techniques to create a system that categorizes accidents into three severity levels; minor, moderate and severe. By leveraging learning capabilities, our method boosts the precision and efficiency of safety protocols in cities. The outcomes exhibit promising results in categorizing accident severity offering a tool for enhancing urban safety infrastructure. Through empowering cities to handle accidents, our model establishes a foundation for safety initiatives. In essence, this study contributes to enhancing safety standards in cities promoting resilience and sustainability, within settings.

Keywords : Smart Cities, Deep Learning, Long Short-Term Memory (LSTM), Accident Classification, Severity Prediction, Urban Safety Infrastructure

I. INTRODUCTION

Global issue in road accidents continues to take lives of many, cause loss to the economy, and result in severe damage to infrastructure. The inclusion of deep learning methods such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) into these systems, however, seems like a feasible and promising approach for enhanced accuracy in prediction due to the rapid development of accident detection technologies. Despite these accomplishments, some issues raised by recent studies are significant challenges that have not yet been addressed [1][2]. For example, the lack of comparison between accident detection techniques with machine learning approaches is one key drawback. Therefore, this study will present a comprehensive review of these emerging trends in road accident severity detection. Road accidents continue to be an international issue, leading to a large number of casualties, losses in the economy, and severe destruction of infrastructure. With the rapid growth in accident detection systems, the incorporation of deep learning methods into these frameworks, such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), has appeared as a promising approach towards improved accuracy in predicting [1][3][9]. Nonetheless, there are difficulties being encountered by recent literature that hinder a comprehensive understanding, such as no investigation into the effectiveness among different accident detection technologies and machine learning models. Given these gaps, this undertaking will aim to do an in-depth review, comparative analysis, and fusion of the emerging technologies for road accident severity detection.

An exhaustive literature review on road accident detection systems, looking into how they can help reduce the level of damage caused by accidents, is described in [1]. The paper examines different technologies such as smartphones, Vehicular Ad-Hoc networks, and GPS. Particular attention is given to the synthesis of Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) models that have many useful applications in practice. Machine Learning Techniques for Crash Injury Severity was proposed by Durgesh Kumar Yadav [1].

Mubariz Manzoor et al. [4] uses Random Forest to uncover important factors strongly associated with highway accident severity. In this paper, the authors propose the integration of machine learning and deep learning models, specifically combining Random Forest with Convolutional Neural Network (RFCNN), as a composite method for predicting road accident severity. The article underlines that traffic accidents can be attributed to many factors.

The system in paper [15] reviewed by Faisal & Almilad uses a microcontroller, GPS, and various sensors to determine physical parameters related to vehicle movement. The study examined various machine learning classifiers and identified Gaussian Mixture Model (GMM) and Classification and Regression Trees (CART) as effective models. Classification of vulnerable road users based on machine learning and also focuses on developing an IoT-based accident detection and classification system. As stated by Awan, Saba & Mehmood, Zahid [10] utilizes deep learning techniques to advance accident se verity modeling in the field of traffic safety. The ASLP

DL framework includes deep neural network (DNN), d eep convolutional neural network and deep recurrent ne ural network. The study claims to outperform traditiona l methods as well as existing machine learning and dee p learning techniques.Accident prediction based on dee p learning.

The research conducted by the author Ayoub Esswidi [6] describes the road safety impact on various sectors and concludes that the use of exploratory data analysis (EDA) and machine learning techniques is necessary for predicting the accident severity of traffic road conditions. This work highlights the capability of artificial neural networks (ANN).

In article [9], the application of Recurrent Neural Networks (RNN) to determine the injury severity level of car accidents is described with an emphasis on RNN's ability in dealing with sequential data. The analysis also shows some disadvantages that classical NN have when they are applied in this kind of problem field have been compared by , Maher & Pradhan, Biswajeet [9].

An issue worthy of attention is the application of deep learning algorithms to predict the severity level of injuries that can occur due to road accidents in Malaysia based on information provided by Feras Chikh Oughali [3]. This study aims to explore the capabilities of Feedforward Neural Networks (NN), Recurrent Neural Networks (RNN), and Convolutional Neural Networks (CNN) in learning temporal and spatial dependencies.

The traffic accident problem stated by M, Girija & V in [11] has not escaped attention worldwide, and this paper offers a new approach using deep learning for forecasting accidents. They gather a large set of data where details like weather conditions, type of road, and prior history of accidents are contained, and run an analysis on that data. Subsequently, they describe in detail Convolutional Neural Networks (CNNs), which play an important role in the potential use to promote road safety.

In the "Crash Injury Severity Analysis" study, the authors Syed Maisur Rahman[2] explore the drawbacks of conventional statistical models and their inability to accurately predict injury severity. The novel machine learning methodologies discussed here employ multi-layer perceptron (MLP) frameworks integrated with distinct embedding layers as well as TabNet. The main objective of this paper is to shed light on how the road safety analysis domain can surpass impediments in statistical models, thereby indicating that innovative methods are at hand.

II. PROPOSED METHODOLOGY

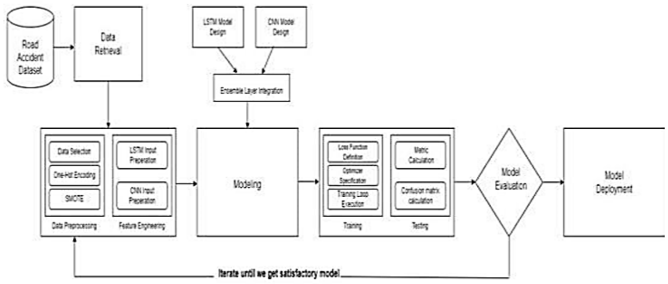

Figure 1. Overview of the Model

The proposed model explores the use of deep learning in enhancing the smart city safety sector. Our strategy involves an ensemble model that incorporates spatial and temporal feature extraction to create a reliable road accident severity detection system. The initial phase is data pre-processing, and it includes cleaning data as well as filling missing values or addressing imbalances between classes. After this stage, we consider the CNN as the most effective feature extractor for spatial features and the LSTM to accurately capture the temporal dependencies. However, it is necessary to note that our ensemble model must combine both techniques such as averaging since these techniques refine severity predictions. Furthermore, the effectiveness of the model was demonstrated by testing based on accuracy, precision, among other metrics that should form part of rigorous model evaluation.

The methodology concludes with iteratively improving the model, which is done by tuning the hyper parameters and architectural design. The solution not only helps in improving smart city safety but also lays down the foundation of a responsive and intelligent urban safety system.

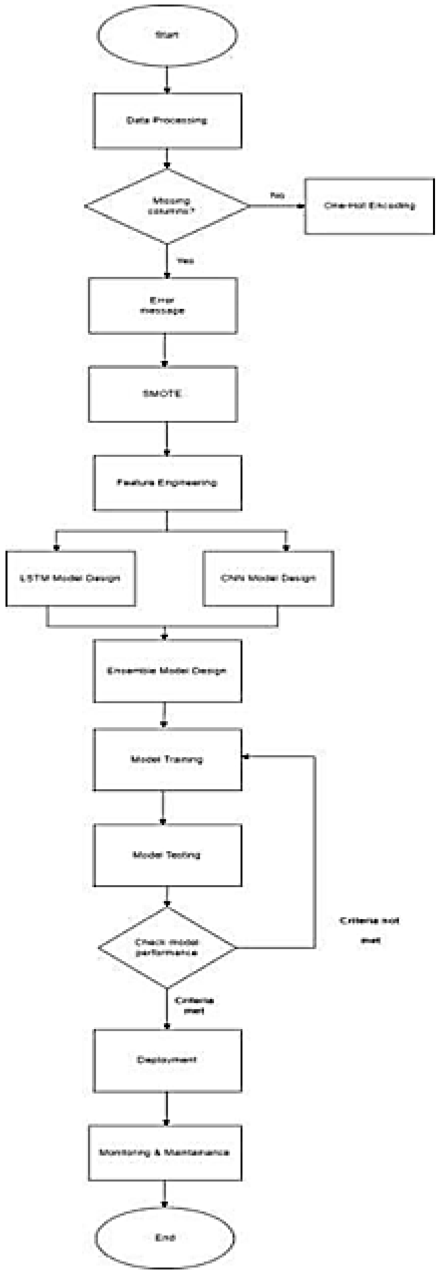

Figure 2. Flow Diagram

A. Dataset Description

The source of information for this study was the data available on Kaggle, which has a multitude of attributes regarding smart cities' traffic crashes. These properties include geographic coordinates, temporal data, road conditions, weather elements, and, more importantly, other factors that affect accidents. The resulting variable is accident severity, which helps in creating an efficient prediction model through a complex deep learning approach.

On the other hand, broad dataset usage, as in this case, helps to view the problem from many sides, thus allowing the obtainment of valuable information about factors that influence accident severity in smart cities. The nature of the data, covering spatial, temporal, and environmental aspects, contributes positively to implementing the program in city centers to enhance safety measures.

Since it can be used for creating and validating an accident classification model, leveraging the dataset is an excellent idea that will positively affect the development of proactive safety measures and contribute to the establishment of intelligent urban systems.

B. Data Pre-Processing

1) Dealing with Missing Values and Outliers: The absence of values is a major machine learning challenge because it may result in the loss of a great amount of information that can directly affect the models. The snippet below exemplifies how the Python code using the Pandas library helps to fill missing numeric values with the mean average from a given column. If you choose this approach, you will keep your data integrity, since none of the data sets contain a gap after any type of change, even an absence, which means they were all filled with some value that provided no loss of information quality.

2) Data Encoding and Normalization: One of the categorical variables, accident severity, has to be changed into a numerical form, which allows machine learning algorithms to use it. In this situation, One-Hot Encoding will be employed, which is a particular type of representation method by which binary vectors are derived from the categorical variable. By using such an encoding, we ensure that the algorithm does not overlook any features and that they play a role in future calculations.

Given the necessity of normalization lies in its ability to put all the features on the same scale so that none will outweigh the other during training. This min-max scaling and z-score normalization code block demonstrates ways to normalize continuous data. While min-max scaling maps onto a specific range (0-1) often, z-score normalization standardizes by subtracting the mean and dividing by the standard deviation.

3) Feature Selection: Feature selection is highly important in road accident severity detection preprocessing and extraction of critical model training

information. Reaching for suitable variables while rejecting useless ones are parts of feature selection.

Convolutional layers in CNNs look for local patterns like edges, while pooling layers help to reduce dimensionality, capturing global features. Thus, CNNs are capable of extracting relevant spatial features from accident data.

LSTM networks specialize in catching temporal dependencies in road accident data that include time of day, traffic flow dynamics and historical trends. Through LSTM layers it is possible to learn long term dependencies which results into a complete understanding about the sequential nature surrounding accidents.

C. Model Building

1) CNN: Convolutional Neural Networks (CNNs) play a role, in extracting features from photos of accident scenes in our project. The setup involves an input layer

defined by the number of time steps layers with ReLU activation max pooling to reduce dimensions and dense layers with ReLU activation. The last layer utilizes softmax activation for class classification highlighting the importance of CNN in capturing complex spatial patterns within accident scenes. This model significantly improves predictions, on accident severity by considering attributes.

2) LSTM: In our study Long Short Term Memory (LSTM) networks are utilized to capture the time related connections, in a series of accident data. The LSTM model consists of LSTM blocks with specified units a connection that skips from the block to the one and Batch Normalization for maintaining stability. By aggregating outputs and implementing layers with ReLU activation the final dense layer employs softmax activation for class categorization. This LSTM structure excels in recognizing and leveraging trends playing a role, in our objective of comprehending and forecasting accident severity based on temporal characteristics.

3) Ensemble Model: We use a blended model (prototypical averaging, weighted averaging or additional intricate algorithms were mentioned). To do this, we import TensorFlow and scikit-learn.As such, the accident's severity prediction process is enhanced by taking into account both spatial and temporal factors of accidents

In general effect it indicates diverse causes that lead to increase or decrease in severe accidents occurrences.

D. Training and Testing Model

Model training is the most crucial stage of model training where the preprocessed data helps the accident severity detection model to learn. The neural network architecture that combines CNNs and LSTMs is fed with the preprocessed dataset and then the weights of this network's neurons are adjusted so that predicted severity labels differ less from actual ones. Such fine tuning is performed via backpropagation as well as optimization algorithms.Subsequent to training, it would be necessary to assess how well the model performs on a different testing set before it can be used for making generalizations about unseen data. In TensorFlow's Keras evaluate method, one can compute such key metrics as loss and accuracy on a test set. This stage plays an important role in determining whether or not overfitting has taken place because, at times, models may perform well when dealing with training data but therefore fail to be generalized on new instances.

E. Performance Evaluation

When checking how well a model is working, there are some key ways to break things down. First off, there's overall accuracy - what percentage of predictions were right out of all the total predictions made.

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

TP = True Positive TN = True Negative
FP = False Positive FN = False Negative

Then you have got precision and recall, which dig deeper into how the model is handling the positive cases(where it predicts something is true). Precision tells you out of all the positive predictions how many were actually correct and recall (also called sensitivity or truepositive rate) looks at out of all the actual positive cases, how many did the model correctly catch.

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

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

You can also use something called an F1 score to balance precision and recall into one metric. It takes theaverage of the two, weighted so that both matter.

$$F1 = \frac{2 * Precision * Recall}{Precision + Recall}$$

Finally, one last way to evaluate performance is to use a confusion matrix. This gives you four categories:

- True positive - model predicted positive and was right
- True negative - model predicted negative and was right
- False positive - model predicted positive but was wrong
- False negative - model predicted negative but was wrong

Having all these different evaluation approaches gives a very thorough picture of where a model is generally effective as well as where it has some blind spots. They highlight different aspects of performance to piece together. well our model is a performing, and more importantly what are the training dynamics looks like. Model Accuracy diagram — As name suggests, % of correct prediction over total instances. Gives a overall idea about how model is doing correctness wise over training and testing phase. Higher better. Increasing curve would show that model is indeed learning, and generalizing the knowledge. Model Loss diagram — It tells us how well model is converging. It tells, how loss function is decreasing over training epochs. Here, the loss function is a numeric value that indicates how well the model are predicting during training. It decreases over time, showing that, telling us that the model is learning.

III. RESULTS AND DISCUSSION

1) Performance Metrics: The model has high accuracy, precision, recall, F1 score on both training and testing datasets which mean that it effectively learns patterns from pre-processed accident data and generalizes well in new cases.

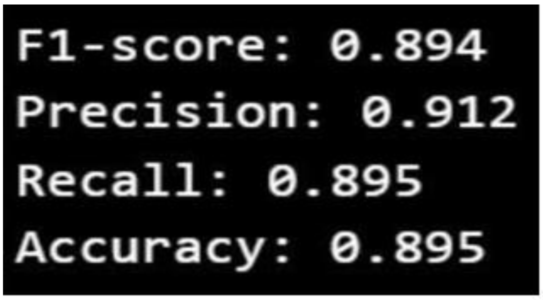
F1-score: 0.894
Precision: 0.912
Recall: 0.895
Accuracy: 0.895

Figure 3. Performance Metrics

2) Confusion Matrix Analysis: Confusion matrix breaks down the model's predictions, showing its ability to correctly predict severe and non-severe accidents. This analysis helps to understand the strengths and improvements areas of the model.

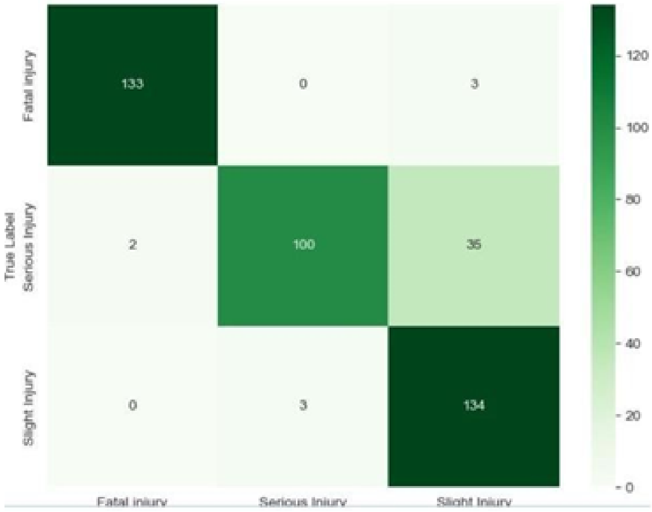

Figure 4. Confusion Matrix Analysis

3) Model Loss and accuracy: Model Accuracy and Model Loss diagrams are the outputs to assess how

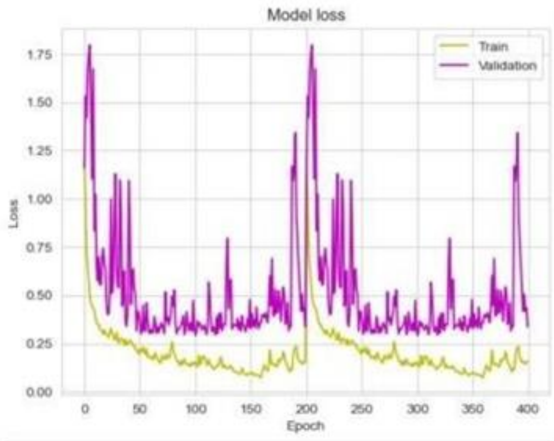

Figure 5. Model Loss

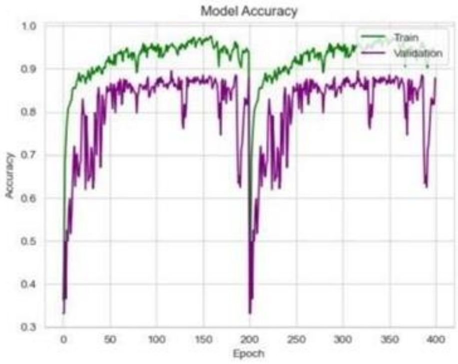

Figure 6 Model Accuracy

4) Future improvements: Continuous monitoring as well as updating models regularly will ensure that this model adapts to changing trends in accident data collection over time. The dynamic model update mechanism discussed in the project timeline would address this aspect explicitly.

IV. CONCLUSION

A milestone in the quest to improve the efficiency of road safety measures is found in the use of deep learning techniques to predict the severity of a road accident. The method uses a combination of CNN and LSTM networks and provides accurate predictions on whether an incident should be categorized as mild, moderate, or severe. Moreover, it can be deduced from this study that the elaborated methodology, starting with data pre- processing using ensemble methods during the modelling stage and going through proper model training and evaluation, demonstrates how rigorously researchers were committed to this research task.

The accurate assessment of diverse levels of accidents in terms of accuracy, precision, recall, and F1 score provides us with evidence of how effectively the model is predicting such cases. It should be noted that this area is currently under serious development; there is a significant demand for evolving a smart city traffic collision detection system employing artificial intelligence solutions as more secure than the current ones and playing a crucial role in ensuring the accomplishment of road safety initiatives

V. REFERENCES

[1]. Renu, Durgesh Kumar Yadav, Iftisham Anjum, Ankita, "Accident Detection Using Deep Learning: A Brief Survey," International Journal of Electronics Communication and Computer Engineering, Vol 11, Issue 3, ISSN

[2]. Karim Sattar, Feras Chikh Oughali, Khaled Assi, Nedal Ratrout, Arshad Jamal, Syed Maisur Rahman, "Transparent deep machine learning framework for predicting traffic crash severity," Springer-Verlag London Ltd, part of Springer Nature 2022

[3]. Maher Al-Zuhairi, Helmi Shafri, Biswajeet Pradhan, Hussain Hamid "Applications of Deep Learning in Severity Prediction of Traffic Accidents,"

[4]. Mubariz Manzoor, Muhammad Umer, Saima Sadiq, Abid Ishaq, Saleem Ullah, Hamza Ahmad Madni, Carmen Bisogni "RFCNN:Traffic Accident Severity Prediction based on Decision Level Fusion of Machine and Deep Learning Model," unpublished.

[5]. Teja Santosh Dandibhotla, "Road Accident Severity Prediction Using Machine Learning Algorithm," J. Article in International Journal of Computer Science Engineering in Research Trends, November 2022.

[6]. Ayoub Esswidi, "Severity Prediction For Traffic Road Accidents," Journal of Theoretical and Applied Information Technology 101.

[7]. Ali Alqatawna, "Highway Traffic Accidents Prediction Using Artifical Neural Network:A Case Study of Freeways in Spain," Journal of Engineering Research Volume 4 (no 1).

[8]. Prudvi Sai Saran Ponduru, "Road Accident Prediction Using LSTM GRU Neural Networks," Thesis submitted, December 2022. [9] Maher Ibrahim Sameen, Biswajeet Pradhan, "Severity Prediction of Traffic Accidents with Recurrent Neural Networks,"

[9]. Al-Zuhairi, Maher & Pradhan, Biswajeet. (2017). Severity Prediction of Traffic Accidents with Recurrent Neural Networks. Applied Sciences. 7. 10.3390/app7060476.

[10]. Awan, Saba & Mehmood, Zahid. (2024). ASLP-DL —A Novel Approach Employing Lightweight Deep Learning Framework for Optimizing Accident Severity Level Prediction. Computers Materials & Continua. 1-10. 10.32604/cmc.2024.047337.

[11]. M, Girija & V, Divya. (2024). Deep Learning-Based Traffic Accident Prediction: An Investigative Study for Enhanced Road Safety. EAI Endorsed Transactions on Internet of Things. 10. 10.4108/eetiot.5166.

[12]. M., Sobhana & Vemulapalli, Nihitha & Mendu, Gnana & Ginjupalli, Naga & Dodda, Pragathi & Subramanyam, Rayanoothala. (2023). URBAN TRAFFIC CRASH ANALYSIS USING DEEP LEARNING TECHNIQUES. Informatyka, Automatyka, Pomiary w Gospodarce i Ochronie Środowiska. 13. 56-63. 10.35784/iapgos.5350.

[13]. T. Jaspar Vinitha Sundari, J. G. Aswathy and S. Jayakamali, "Accident Detection and Severity Analysis using Deep Learning," 2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA), Coimbatore, India, 2021, pp. 1-5, doi: 10.1109/ICAECA52838.2021.9675657.

[14]. T. Jaspar Vinitha Sundari, J. G. Aswathy and S. Jayakamali, "Accident Detection and Severity Analysis using Deep Learning," 2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA), Coimbatore, India, 2021, pp. 1-5, doi: 10.1109/ICAECA52838.2021.9675657.

[15]. Balfaqih, Mohammed & Alharbi, Soltan & Alzain, Moutaz & Alqurashi, Faisal & Almilad, Saif. (2021). An Accident Detection and Classification System Using Internet of Things and Machine Learning towards Smart City. Sustainability. 14. 210. 10.3390/su14010210.

[16]. Mckenna, H. Patricia. (2021). Learning and Data in Smart Cities. 10.1007/978-3-030-70821-4_5.

[17]. Ayman, Shehab Eldeen & Hussein, Walid & Karam, Omar. (2021). Accident detection and road monitoring in real time using deep learning and lane detection algorithms. 30-37. 10.1145/3490725.3490730.

[18]. Komol, Md. Mostafizur & Hasan, Md.Mahmudul & Elhenawy, Mohammed & Yasmin, Shamsunnahar & Masoud, Mahmoud & Rakotonirainy, A.. (2021). Crash Severity Analysis for Queensland Vulnerable Road Users using Machine Learning (Conference abstract).

[19]. Shamenna, Aklilu & Bezabih, Anteneh & Bedane, Ashenafi. (2021). Spatial Distribution of Road Traffic Accident… Spatial Distribution of Road Traffic Accident at Hawassa City Administration, Ethiopia. Ethiopian Journal of Health Sciences. 31. 793-806. 10.4314/ejhs.v31i4.14.

[20]. Ghasedi, Meisam & Sarfjoo Kasmaei, Maryam & Bargegol, Iraj. (2021). Prediction and Analysis of the Severity and Number of Suburban Accidents Using Logit Model, Factor Analysis and Machine Learning: A case study in a developing country. SN Applied Sciences. 3. 10.1007/s42452-020-04081-3.