Contents lists available at SCCE

Journal of Soft Computing in Civil Engineering

Journal homepage: www.jsoftcivil.com

Forecasting Road Accidents Using Deep Learning Approach: Policies to Improve Road Safety

Vijay S. Angadi [1]; Shivaraj Halyal [2*]

1. School of Civil Engineering, KLE Technological University, India
2. Ph.D., School of Civil Engineering, KLE Technological University, India
Corresponding author: *shivaraj.halyal@kletech.ac.in*

https://doi.org/10.22115/SCCE.2023.399598.1654

ARTICLE INFO

Article history:
Received: 29 May 2023
Revised: 08 August 2023
Accepted: 31 December 2023

Keywords:
Accident;
Time Series Forecasting (TSF);
Deep Learning (DL);
Policy-making.

ABSTRACT

The development of smart cities holds immense significance in shaping a nation's urban fabric and effectively addressing urban challenges that profoundly impact the economy. Among these challenges, road accidents pose a significant obstacle to urban progress, affecting lives, supply chain efficiency, and socioeconomic well-being. To address this issue effectively, accurate forecasting of road accidents is crucial for policy formulation and enhancing safety measures. Time series forecasting of road accidents provides invaluable insights for devising strategies, enabling swift actions in the short term to reduce accident rates, and informing well-informed road design and safety management policies for the long term, including the implementation of flyovers, and the enhancement of road quality to withstand all weather conditions. Deep Learning's exceptional pattern recognition capabilities have made it a favored approach for accident forecasting. The study comprehensively evaluates deep learning models, such as RNN, LSTM, CNN+LSTM, GRU, Transformer, and MLP, using a ten-year dataset from the esteemed Smart Road Accident Database in Hubballi-Dharwad. The findings unequivocally underscore LSTM's superiority, exhibiting lower errors in both yearly (RMSE: 0.291, MAE: 0.271, MAPE: 6.674%) and monthly (RMSE: 0.186, MAE: 0.176, MAPE: 5.850%) variations. Based on these compelling findings, the study provides strategic recommendations to urban development authorities, emphasizing comprehensive policy frameworks encompassing short-term and long-term measures to reduce accident rates alongside meticulous safety measures and infrastructure planning. By leveraging insights from deep learning models, urban development authorities can adeptly shape the urban landscape, fostering safer environments and contributing to global safety and prosperity.

How to cite this article: Angadi VS, Halyal S. Forecasting road accidents using deep learning approach: policies to improve road safety. J Soft Comput Civ Eng 2024;8(4):27–53. https://doi.org/10.22115/scce.2023.399598.1654

1. Introduction

Transportation engineering is a multidisciplinary field that harnesses the power of technology and scientific knowledge to design, operate, and manage transportation systems. These systems encompass various modes such as railways, roads, air, and water transit, and their efficient functioning profoundly impacts the economy, particularly in terms of logistics and trade facilitation. Among the modes, railways are a cost-effective means for domestic goods transport, while sea routes offer affordable options for international trade. Over time, roadways have served as a crucial backbone of transportation, witnessing significant advancements in planning, design, construction, and maintenance. However, road accidents continue to pose a substantial challenge for the road transport system despite these improvements.

India, in particular, has grappled with the grim reality of road accidents. In 2021 alone, a staggering 412,432 road accidents were reported, leading to a loss of 153,972 lives and injuries to 384,448 individuals. More alarmingly, the age group most affected by these accidents was between 18 to 45 years, accounting for approximately 67 percent of the total fatalities. These incidents have far-reaching implications for the nation's progress, resulting in the loss of prime-age adults from the workforce and diminishing productivity due to the incapacitating nature of injuries. The complex nature of mixed traffic significantly contributes to accidents [1].

Addressing this critical issue necessitates a comprehensive approach in the form of road safety plans. These plans must thoroughly evaluate injuries, fatalities, and accident data to gain insights into the current situation and assess the efficacy of existing safety measures. To ensure an effective strategy, planners must carefully scrutinize regulations and precautions, considering future events and taking proactive actions [2–5].

Policymaking is pivotal in shaping various aspects of the transportation sector, aiming to advance safety, infrastructure, economics, planning, and management. These strategic policies are crafted based on data derived from diverse processes, with forecasting playing a prominent role. Forecasted accident data proves vital for safety planners as it aids in evaluating existing safety measures, such as infrastructure quality and legislation. Armed with these insights, detailed discussions can be conducted, and effective policies formulated to minimize accident rates and enhance safety in specific geographical areas.

Different methods are employed in accident prediction, including statistical, machine learning, and deep learning techniques. Among these, machine learning stands out for its model tolerance, requiring minimal assumptions, and its computational flexibility, which yields superior performance [6]. Time series forecasting, a subset of machine learning, relies on artificial neural networks, enabling faster and more accessible results.

In pursuit of reducing accident rates, two types of forecasting come into play. Short-term forecasting focuses on the current road accident situations, leading to the formulation of short-term policies such as road repairs and speed regulations. Typically, this forecasting covers a period of 1 to 5 years. On the other hand, long-term forecasting delves into analyzing road accident scenarios to develop comprehensive, enduring policies, including new road planning and expansion. This forecasting is usually conducted for a duration of 15 to 30 years.

The application of deep learning to predict road accidents in developing countries faces limitations, as most studies either rely on a single model or compare just two models. This approach generates uncertainty about their efficiency and fails to harness the full potential of deep learning. Moreover, the intricate traffic environments in these nations, driven by rapid infrastructure projects that lead to intermittent road blockages and frequent crashes, necessitate attention. Often, spatial and temporal dimensions of road accidents are disregarded. To address these concerns, this study centers on forecasting road accidents in the urban regions of Hubballi-Dharwad by comparing diverse efficient deep-learning models known for their exceptional pattern recognition capabilities. A comprehensive literature review meticulously selects appropriate models, followed by assessing their performance using RMSE, MAE, and MAPE metrics with training and testing data. The outcomes yield crucial insights into enhancing road safety, shaping safety protocols, and infrastructure planning. This research establishes itself as a benchmark for accident forecasting studies in developing countries, adding to the global drive for safer roads.

2. Literature review

The literature review helps the researcher know the advanced solutions and activities done by other researchers worldwide to solve their research problem. A thorough literature review was conducted to find and explore the views and perspectives of other researchers on the subject.

2.1. Accidents

An accident is an unplanned and uncontrollable incident in which a person or object's activity or reaction causes human injury or property damage. A traffic accident can be attributed to the road-vehicle-driver system's incapacity to fulfill necessary tasks for a safe trip, potentially causing harm. The predominant factors include the road network's insufficient maintenance, ineffective systematic enforcement, and the resultant absence of safety measures, all contributing to accidents. [7].

The effects of traffic accidents on business and human health are significant. The following have been identified as the primary contributing factors to accidents: speed, alcohol, seat, helmet, child seat, vision, driving impairment, weariness, adverse effects on children and youth on the road, and epidemics [8].

2.2. Forecasting techniques

Anticipating accidents stands vital for road safety, entailing proactive steps to prevent their occurrence through precautionary measures [9].

2.2.1. Statistical techniques

Jha carried out comparison research on time analysis-based traffic forecasting. Regression modeling was used to determine the cause of auto accidents. Compare the outcomes to those discovered through temporal analysis. The comparison demonstrates that time series are more accurate and efficient at forecasting [10].

In Zimbabwe, Mutangi examined road accidents to create a forecasting model for future occurrences. The study assessed the model's performance using diverse methods and concluded that ARIMA(0,1,0) accurately predicted the annual traffic accident rate [11].

MA Quaddus used the ARIMA model to evaluate data on traffic accidents in the UK. The authors used statistical value autoregressive (INAR), Poisson, and negative binomial (NB) approaches for their investigation. This study demonstrates that the model fit of ARIMA and INAR Poisson is equivalent when using the cumulative data series of traffic accidents. The authors conclude that if the data are presented as time-disaggregated, the ARIMA model does not follow the INAR Poisson model [12].

Numerous research endeavors have tackled the serial dependence challenges within ARIMA through statistical means. Diverse teams have presented models aimed at enhancing outcomes, including the DRAG model—a specific ARIMA instance—and the Auto-Regressive (AR) model tailored for road safety assessment, among other innovative approaches [13].

ARIMA and SARIMA models were utilized for Short-term passenger demand modeling using automatic fare collection data in Hubballi-Dharwad Bus Rapid Transit System (HDBRTS). The study showed that SARIMA outperformed ARIMA based on MAPE [14]

Rabbani and co-researchers focused on Peshawar, Pakistan, utilizing SARIMA and ES techniques for accident rate prediction and temporal pattern identification. Through assessment of vital metrics—mean absolute error (MAE), root mean square error (RMSE), mean absolute percentage error (MAPE), and normalized Bayesian information criterion (BIC)—the superiority of the ES model over SARIMA in accident data prediction was established. The ES (Simple seasonal) model's remarkable accuracy was evident, boasting the lowest values—MAE: 1.71, RMSE: 2.20, MAPE: 17.73, and BIC: 1.65. When quantitatively aligning the models with actual 2019 accident data, SARIMA exhibited two errors, whereas the ES model displayed just one, closely matching the death count. As such, the ES model displayed enhanced performance in reflecting observed accident counts, confirming its superiority over the SARIMA model [15].

Caliendo employed the Poisson, Negative Binomial, and Negative Multinomial regression models to predict the number of accidents in multi-lane roadways and found that they were helpful [16].

The study aims to use Bayesian reasoning to estimate the number of accidents. The findings indicate that considering the link between accident counts and overdispersion might lead to more accurate estimations. According to the results of this investigation, the Poisson-Weibull model can predict more accidents. Traffic volumes, access, open areas, and schools indicate the likelihood of an accident. This safety research may aid in the improvement of road safety rules [17,18].

2.2.2. Machine and deep learning

Recurrent Neural Networks (RNN) have been the pioneer in the field of time series forecasting due to their advantage of the artificial neural network concept, which utilizes the speed and efficiency of neurons in humans [19–21], LSTM (Long Short Term Memory), an upgraded version of RNN, has replaced RNN as the popular choice for time series forecasting [22–26].

The combination of CNN with LSTM, known as CNN+LSTM, has proved effective for time series forecasting as the output of CNN layers is considered input for LSTM. The efficiency of the hybrid neural network model is high and has applications in forecasting the gold price, stock price, energy consumption, crash prediction, etc [27–35].

To tackle spatial heterogeneity, researchers introduced the Hetero-ConvLSTM framework, incorporating spatial graph features and model ensemble. Testing on 8 years of Iowa's state-wide data demonstrated the superiority of the Hetero-ConvLSTM over baselines in prediction accuracy [36].

A novel Traffic Accident Severity Prediction-Convolutional Neural Network (TASP-CNN) model with traffic accident integration is presented for severe traffic accidents. An image attribute matrix to grayscale (FM2GI) technique is developed, which transforms the relationship between traffic accident data into a grayscale picture with linkages formed according to different notions. pattern. Experiments also reveal that the hint that the forecast has a heavy collision performs better [37].

The long short-term memory, gradient-boosted regression tree (LSTM-GBRT), was used to create the traffic accident prediction model, and the accident-related data is used to estimate the safety indicators of traffic accidents. Comparing the LSTM-GBRT model to several transformations and neural network models, experimental findings demonstrate that it produces accurate and reliable outcomes. Traffic control centers can better grasp the state of traffic safety thanks to the LSTM-GBRT model's ability to anticipate the safety of traffic accidents [38].

A novel deep spatial-temporal graph convolutional network (DSTGCN) is presented to forecast traffic accidents. General world data were gathered to assess the suggested model, including accident statistics, urban traffic speed, road network, weather prediction, and areas of interest. Experimental findings on real-world data demonstrate that DSTGCN outperforms traditional approaches and cutting-edge techniques [39].

GRU (Gated Recurrent Unit) is an upgraded version of LSTM. It has been observed to be easier to train and quicker to obtain results, which has helped in various applications like time series forecasting, image recognition, etc [40–43].

Transformer is a newer generation deep learning model developed to cater to the limitations of RNN, LSTM, and GRU and is steadily gaining importance in time series forecasting by removing the vanishing gradient problem and shortening the time to train [30,44].

MLPs (Multi-Layer Perceptrons) are used nowadays in time series forecasting due to their ability to solve complex non-linear problems. They handle large amounts of input data well and make quicker predictions after training [45–47].

Various studies have employed several innovative Machine Learning (ML) and Deep Learning (DL) models alongside classification algorithms to assess road safety and propose improvement strategies. The objective is to combine these models, mitigating their limitations and enhancing their effectiveness in evaluating road and transportation systems' safety. The models used include Artificial Neural Network (ANN) and Grey Wolf Optimization algorithm (GWO)–ANN [48], Stepwise Weight Assessment Ratio Analysis (SWARA) and Measurement of Alternatives and

Ranking according to COmpromise Solution (MARCOS) approach under Spherical Fuzzy (SF) [49], binary model of Neural Networks and the Group Method of Data Handling (GMDH) [50], Particle Swarm Optimization (PSO) and Genetic Algorithm (GA) with the k-means algorithm [51], Group Method of Data Handling (GMDH)-type Neural Network, and a combination of Support Vector Machine (SVM) and the Grasshopper Optimization Algorithm (GOA) [52], as well as Artificial Neural Network (ANN) with Particle Swarm Optimization (PSO) and Differential Evolution (DE) algorithm [53]. The review of Intelligent Transportation Systems (ITS) aims to explore emerging research areas for improving road safety [54].

2.3. Literature insights and objectives of the study

Various modeling approaches have been compared, revealing that statistical models are less efficient than machine and deep learning models. This discrepancy is primarily attributed to the characteristics of the datasets involved. Statistical models are better suited for smaller and univariate data, whereas machine and deep learning models excel with more extensive and multivariate data. They are often combined with boosting algorithms to enhance the reliability and efficiency of statistical models. However, this combination introduces complexity to the study, requiring recurrent training for different models and algorithms, leading to longer durations and making it unsuitable for shorter-term studies. Nonetheless, this approach has shown promise in some instances. Among the deep learning models, GRU and Transformer models have emerged as more efficient than RNN and LSTM models. This advantage can be attributed to their adaptability across diverse fields and easier training processes. Nevertheless, to conclusively determine the superiority of GRU and Transformer models over RNN and LSTM models, a comprehensive study comparing their efficiency in various scenarios using the same dataset is essential. Furthermore, there is a need for further research focused on the temporal variation of accident forecasting.

To address these gaps in the existing research, this study adopts deep learning techniques to forecast road accidents. It compares the temporally varied results to identify the most suitable model with minimal errors. The models evaluated in the research include RNN, LSTM, CNN+LSTM, GRU, Transformer, and MLP. The considered error metrics are RMSE, MAE, and MAPE. Using the results from the selected model with the highest suitability, the study aims to recommend strategies for improving the accident rate in the urban regions of Hubballi-Dharwad. These recommendations encompass short-term and long-term policies involving safety measures and infrastructure planning. The goal is to develop effective measures that can reduce the number of accidents and enhance road safety. In addition to addressing the forecasting aspect, the study investigates the efficiency of GRU and Transformer models compared to RNN and LSTM models. The objective is to determine their applicability in different scenarios and ascertain their performance under various conditions.

In summary, statistical models are better for smaller datasets but less efficient than machine and deep learning models, especially for more extensive and multivariate data. Boosting algorithms enhances their reliability but increases complexity and training times. GRU and Transformer show higher efficiency, but their superiority over RNN and LSTM needs further research. This study uses deep learning to forecast road accidents, comparing results over time to identify the most suitable model. The findings will inform road safety strategies in urban areas, encompassing short-term and long-term policies, safety measures, and infrastructure planning. Additionally, the study explores GRU and Transformer efficiency in various scenarios for broader applicability understanding.

3. Methodology

Figure 1 depicts the research methodology, which begins with a comprehensive literature review to identify the most suitable method for predicting traffic accidents, in this case, deep learning. Data on traffic accidents is then collected and preprocessed to create the necessary dataset. The chosen models are trained with the training data and tested using the real dataset to evaluate their performance and errors. Based on the identified errors, the most effective model is selected. The model's predictions are then utilized to provide recommendations for improving the accident scenario. This flowchart represents the entire investigation in a single sequence summary.

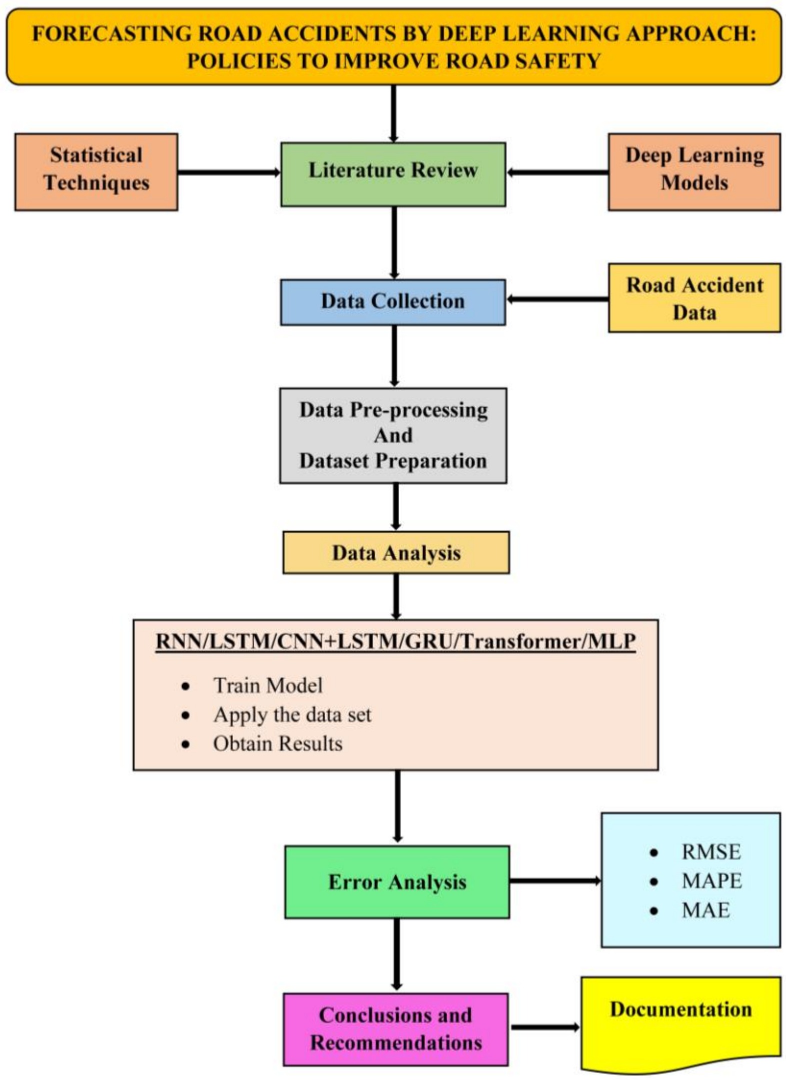

Fig. 1. Proposed Methodology for the Current Investigation.

4. Study area

Located in Karnataka, India, Hubballi and Dharwad are twin cities. Hubballi-Dharwad, the state's second-largest urban agglomeration, comprises Dharwad's administrative center and the commercial hub Hubballi. The cities are governed by the Hubballi-Dharwad Municipal Corporation (HDMC), and it is the state's second-largest municipality in terms of area after Bangalore. Based on the interim data from the 2011 Census, the combined population of the twin cities is 943,788, and they are categorized as urban areas. Hubballi–Dharwad witnessed a population growth of 22.99% from 1981 to 1991 and 21.2% from 1991 to 2001 [55].

The city boasts of being well-connected by air, rail, and road. Hubballi Airport is the second busiest airport in Karnataka, connecting significant destinations like Bengaluru, Mumbai, Chennai, etc. The Golden Quadrilateral connects the cities. Sri Siddharoodha Swamiji Railway Station holds the world record for the most extended railway platform and also acts as a major interchange to many destinations. Figure 2 depicts the urban regions of Hubballi-Dharwad cities, which are considered for the current study.

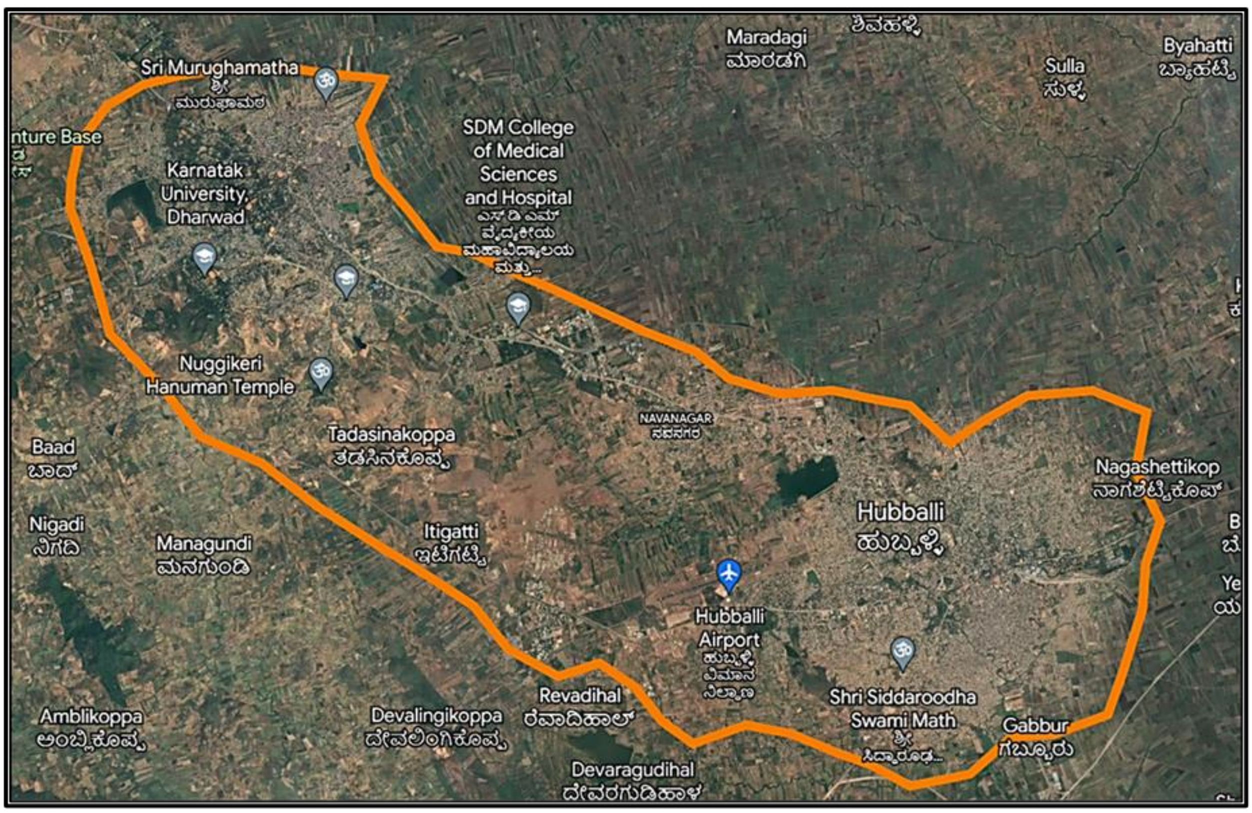

Fig. 2. Urban regions of Hubballi-Dharwad (Credits: Google Earth, 2023).

5. Data source and processing

For this study, the primary source of data is the Smart Road Accident Database, which is meticulously maintained by the Hubballi-Dharwad Police Commissioner's Office. This comprehensive dataset spans a ten-year period (2013 to 2022), concentrating specifically on road accidents within the urban regions of the Hubballi-Dharwad twin cities. The dataset encompasses a wide array of multivariate balanced information crucial for conducting an in-depth analysis of road accidents in the area.

The dataset includes essential accident-related variables, such as accident counts, which provide an overall view of the frequency of accidents over the ten years. Additionally, it contains data on fatalities, minor injuries, and casualties, enabling researchers to gain insights into the severity of these accidents and their impact on the local community.

Moreover, the dataset also includes relevant contextual variables such as weather conditions, road infrastructure, traffic volume, and other contributing factors that can significantly influence road safety. By incorporating these variables, researchers can better understand the complex dynamics of road accidents and identify potential patterns or correlations.

Analyzing such a rich and extensive dataset through deep learning models will enable the research team to make accurate predictions and evaluations. Using a decade-long dataset allows for examining long-term trends and patterns, which can further enhance the study's reliability and applicability.

Overall, the utilization of the Smart Road Accident Database from the Hubballi-Dharwad Police Commissioner's Office provides a robust foundation for this research, enabling comprehensive assessments of road accidents in the urban regions of the twin cities and facilitating the development of effective strategies to improve road safety and reduce accidents.

Data division into training and testing sets is crucial for accurate study results. Researchers have explored different proportions to balance training data volume and test data representativeness. For instance, Halyal et al. successfully employed an 80% training and 20% testing data split to forecast passenger demand, significantly reducing errors [26]. Similarly, other investigations have opted for ratios such as 70:30 [56,57] and 60:40 [58], attaining satisfactory outcomes that align with their objectives.

Drawing from an exhaustive literature review and a critical assessment of the results, our current study adopts an 80:20 data division ratio. This implies that 80% of the data is dedicated to training the predictive models, while the remaining 20% is reserved for robustly evaluating the model's performance during testing. This approach has been strategically chosen to ensure an ample training dataset, fostering the development of accurate and generalized models. Simultaneously, the testing dataset is kept sufficiently comprehensive to guarantee the assessment of the model's predictive capabilities in real-world scenarios.

By adopting this methodological choice, the current research aims to elevate the quality of the study's findings, enhancing the reliability and generalizability of the results. The carefully balanced data split fosters a practical model that can yield valuable insights into passenger demand forecasting, contributing to more informed decision-making and efficient resource allocation in transportation systems.

6. Data analysis

Deep Learning, a branch of Machine Learning, is chosen as the preferred approach. Deep Learning employs neural networks with multiple layers to recognize complex patterns in diverse data types, enabling accurate predictions. The study has used RNN, LSTM, CNN+LSTM, GRU, Transformer, and MLP models for time series forecasting based on prior research.

6.1. RNN

RNNs (Recurrent Neural Networks) are a learning model that recursively updates new state h_t using previous state h_{t-1} and current input x_t. It can also be described as a network of multiple cells connecting in series along the time axis. Each cell in the network computes its state at a specific time step. Figure 3 represents the architecture of RNN.

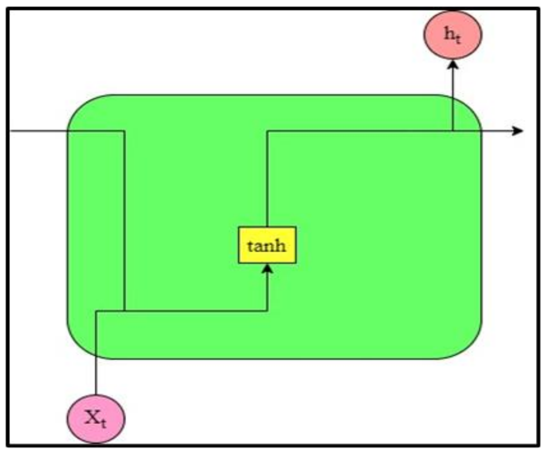

Fig. 3. RNN Architecture.

An RNN cell is represented by Equation 1:

$$h_t = \sigma(Wx_t + Uh_{t-1} + b) \quad (1)$$

In this equation, x_t and h_t are the cell's entry and latent states at time t, respectively. W and U are the current cell's weight matrix and the input loop's weight matrix, and b represents the neuron's bias. $\sigma()$ is the function of the neuron (sigmoid function).

RNNs have a significant drawback, the vanishing gradient problem, caused by the repetitive use of the weight matrix. As the network processes sequences over time, the gradient gradually diminishes, leading to the RNN's inability to retain information for an extended period. Consequently, RNNs need help with long-term dependencies, limiting their ability to effectively capture and utilize relevant context or information from distant time steps. This issue hinders the model's capacity to retain crucial information, making it remember only short-term patterns and impeding its performance in tasks that require long-range temporal dependencies [59].

6.2. LSTM

Long Short-term Memory (LSTM) networks are a variant of RNNs that partially address the vanishing gradient problem and learn longer-term dependencies in the time series data. LSTM uses the concept of a gate in a cell which accounts for different operations such as input, forget, and output. Figure 4 represents the architecture of LSTM.

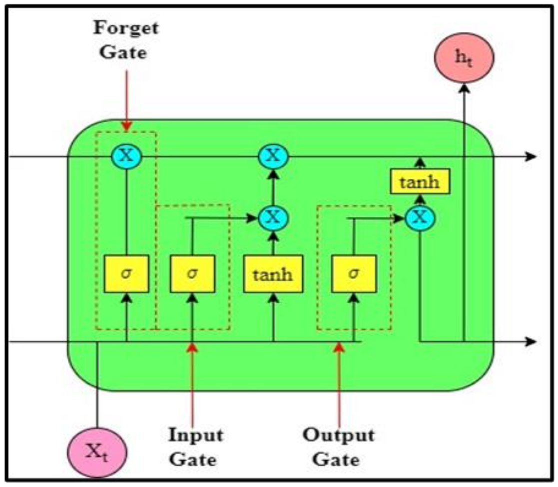

Fig. 4. LSTM Architecture.

The following equations 2 to 7 represent the working of the LSTM cell:

$$f_t = \sigma(W_f \,.x_t + U_f.h_{t-1} + b_f) \tag{2}$$

$$i_t = \sigma(W_i.x_t + U_i.h_{t-1} + b_i) \tag{3}$$

$$\tilde{c}_t = \tanh(W_c.x_t + U_c.h_{t-1} + b_c) \tag{4}$$

$$c_t = f_t.c_{t-1} + i_t.c_t \tag{5}$$

$$o_t = \sigma\,(W_o.x_t + U_o.h_{t-1} + b_o) \tag{6}$$

$$h_t = o_t.\tanh(c_t) \tag{7}$$

Where x_t and h_t are the input and latent states at time t. W_f, W_o, and W_i are the weighted time inputs of the memory, output, and input tables. U_f, U_o, and U_i are the equivalent weights of the memory, output, and input gate from the data cycle at time t, respectively. b is the bias of the neuron. For activation, we have two: the sigmoid activation function for *σ()* neurons and the *tanh()* hyperbolic tangent activation function. c_t represents the cell state of the LSTM at time t.

6.3. CNN+LSTM

Deep learning uses convolutional neural networks (CNNs). It is one of several neural network tools for various tasks and information. Deep learning algorithms use a specific network architecture called CNN to recognize pixels and pictures. CNN is another neural network that can show significant real-time and visual data. As a result, picture-related tasks like classification and pattern recognition depend on image recognition. CNNs employ techniques from linear algebra, such as matrix multiplication, to identify picture patterns. CNNs may also transmit signal and audio data. In-depth learning Convolutional, pooling, and full link (FC) layers are the three levels of CNN. The first layer is the convolution layer, while the last is the FC layer. Figure 5 shows how CNN is organized.

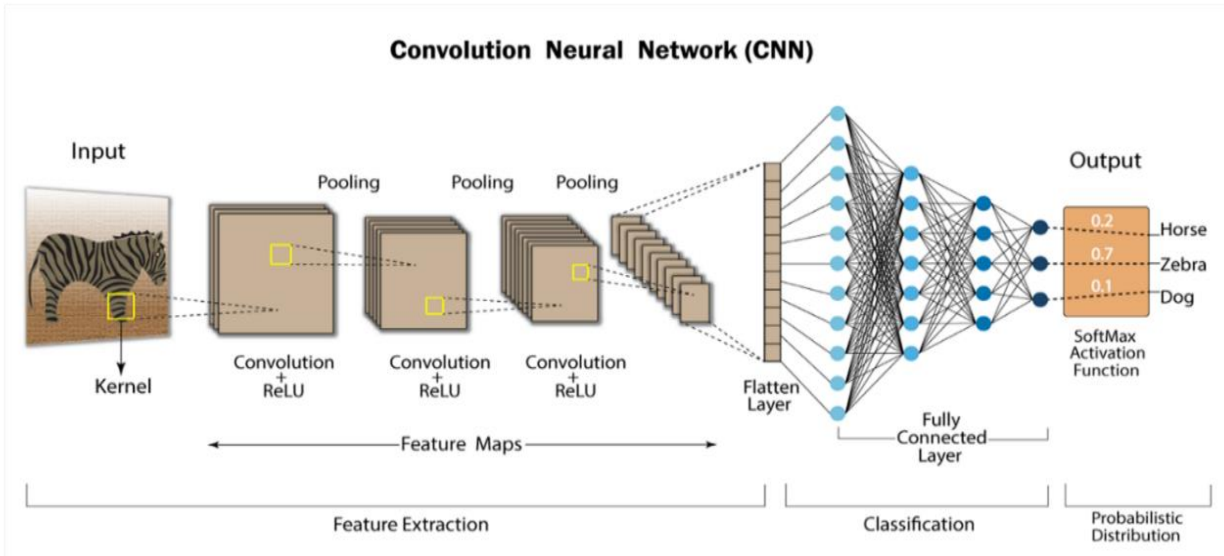

Fig. 5. CNN Architecture.

The proposed CNN-LSTM model merges Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) to process time data effectively. This innovative design adapts the time data to suit CNN's input format by incorporating LSTM's capabilities. The model comprises two key layers: the CNN layer, responsible for extracting crucial features from real-time data, and the LSTM layer, which performs the final estimation. Figure 6 illustrates the architecture of the CNN-LSTM model, showcasing how these integrated layers work together to enhance the model's performance in handling time-based information and producing accurate predictions.

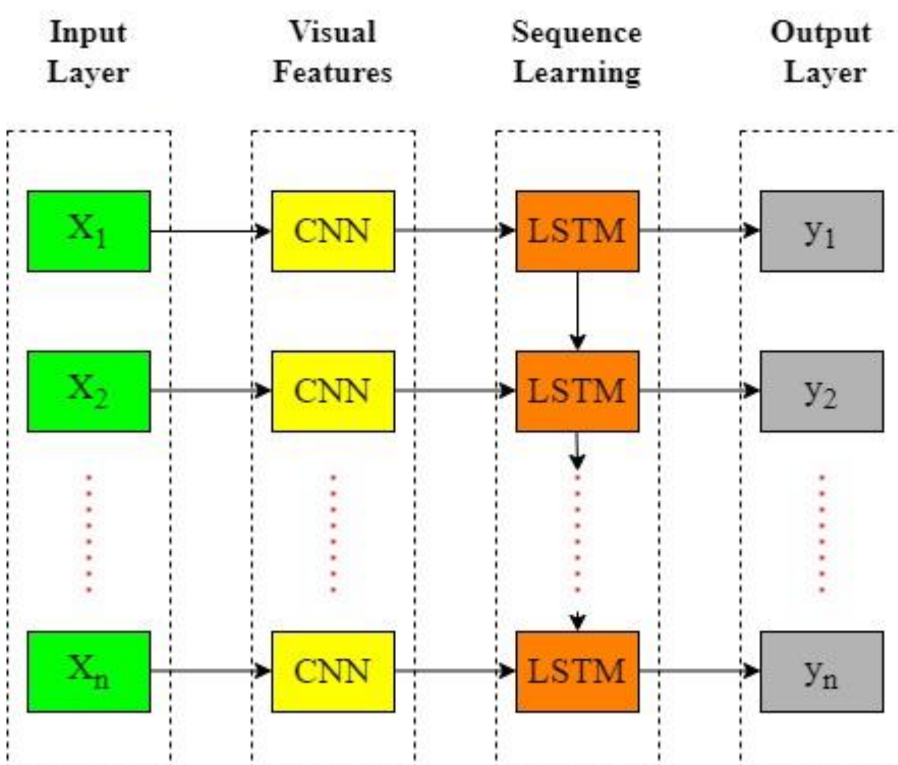

Fig. 6. CNN+LSTM Architecture.

6.4. GRU

GRUs, which stand for Gated Recurrent Units, are a modified version of LSTMs that address inconsistencies by combining the "input" and "forget" gates into a single "update" gate. This update gate represents the cost associated with data storage, while the reset gate represents the cost of forgetfulness. These adjustments, illustrated in Figure 7, optimize the storage and forgetting mechanisms of the GRU unit, ultimately minimizing inconsistencies. By reducing the number of gates, GRUs enhance the efficiency and performance of recurrent neural networks, providing a more streamlined approach to managing information retention and forgetting.

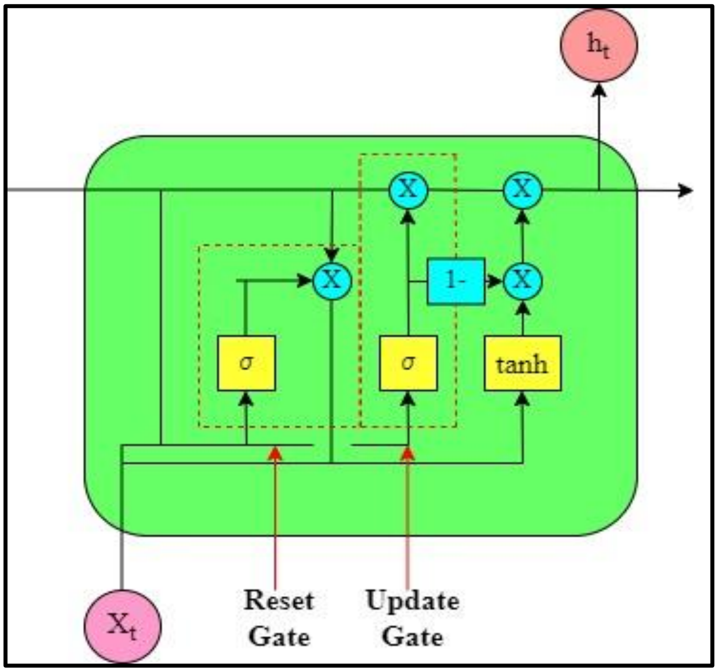

Fig. 7. GRU Architecture.

The working of GRU is expressed in the following equations 8 to 11:

$$z_t = \sigma\ (W_z.x_t +\ U_z.h_{t-1} + b_z) \quad (8)$$

$$r_t = \sigma\ (W_r.x_t +\ U_r\ .h_{t-1} + b_r) \quad (9)$$

$$\tilde{}h_t = \tanh(W_h.x_t + U_h\ .r_t.h_{t-1} + b_h) \quad (10)$$

$$h_t = (1\ -\ z_t).h_{t-1} +\ z_t\ .\tilde{}h_t \quad (11)$$

Where x_t and h_t are the input and latent states at time t, respectively. W_r and W_z are weighted inputs of reset and update gates. U_r and U_z are the equivalent weights of gate reset and gate update from the data cycle at time t. b is the bias of the neuron. For activation, we have two: the sigmoid activation function for *σ()* neurons and the *tanh()* hyperbolic tangent activation function.

6.5. Transformer

Recurrent Neural Networks (RNNs) are purpose-built for processing input sequences of varying lengths, making them highly suitable for sequential data. Unlike conventional neural networks, RNNs establish connections between sequence inputs, facilitating the flow of information among neighboring elements. RNNs find applications in diverse models such as Vector-Sequence, Sequence-Vector, and Sequence-to-Sequence. Nevertheless, RNNs encounter challenges when dealing with long sequences, including slow training and the vanishing gradient problem.

To address the vanishing gradient problem, Long Short-Term Memory (LSTM) was introduced as a specialized variant of RNNs. LSTMs excel at capturing long-term dependencies by incorporating specialized neurons that employ a mechanism allowing information to bypass extensive processing within the current cell. This mechanism enables LSTMs to retain memory over extended periods. However, LSTMs may require longer training times and can be slower than traditional RNNs.

Transformer networks have garnered considerable recognition for their exceptional ability to handle sequences, mainly due to their attention mechanisms. These mechanisms empower the network to assign importance to relevant past information. The Transformer architecture comprises encoder and decoder components. Within the decoder, the Masked Attention and Multi-Head Attention mechanisms effectively leverage information from the encoder's output. The decoder selects a feature vector and, during training, provides future information, influencing the encoder output. The encoder receives input from a dimension-inverted perspective and generates a vector representation the decoder utilizes. Figure 8 represents the mechanism of Transformer architecture.

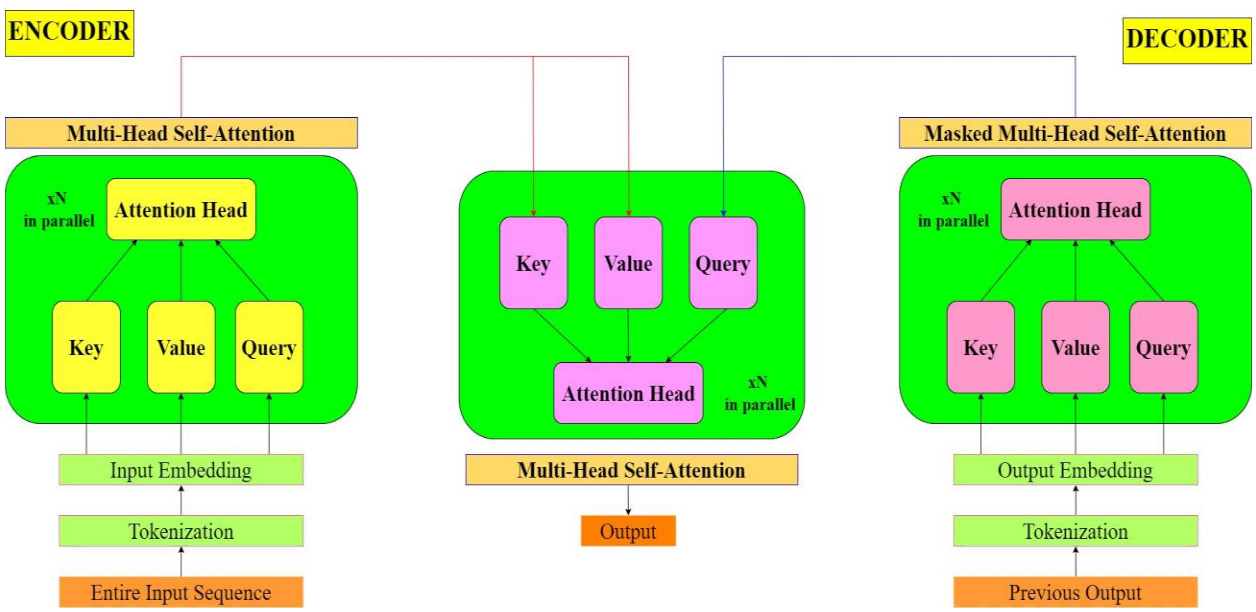

Fig. 8. Transformer Architecture.

6.6. MLP

The term "multilayer perceptron" (MLP) refers to a multilayer neural network. They have a layer of neurons or more. Data is fed into the input process, and predictions are made on the output (also called the visible layer), which may include one or more hidden methods that give some abstraction. The network's initial layer, the input layer, uses the input to inform the code. A second hidden process processes the data from the input process. Figure 9 represents the MLP model architecture.

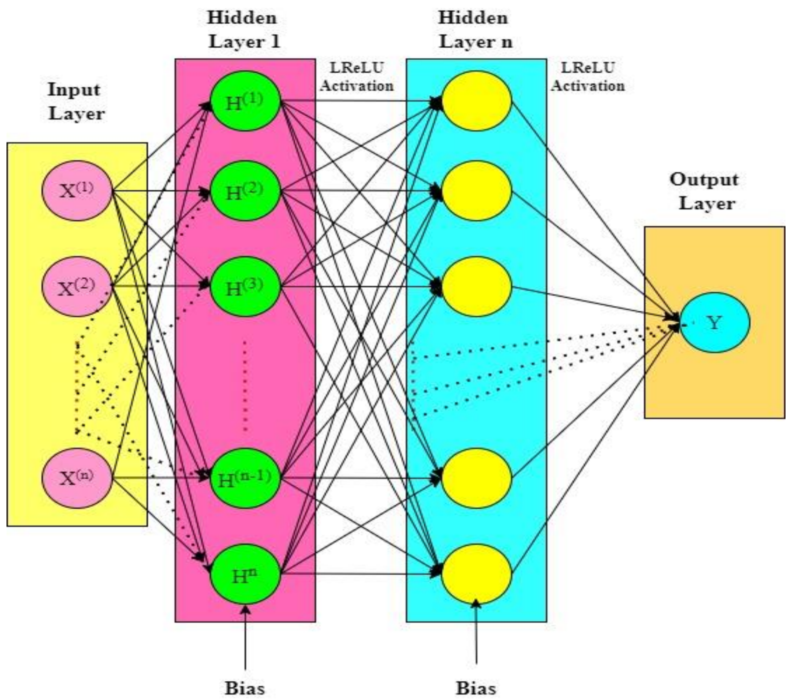

Fig. 9. MLP Architecture.

6.7. Model development process

The model development process plays a crucial role in understanding the inner workings of a model and achieving successful results. In this study, six different models were utilized for road accident forecasting, and the following explanation outlines the model development procedure for LSTM. Similar steps are followed for the other models.

The LSTM model is implemented using the Anaconda Python distribution with Python 3.11 as the programming language. The Jupyter Notebook 7.0 IDE is employed, leveraging the TensorFlow and Keras APIs to construct and train the LSTM model. An AMD Radeon™ 520 Graphics card provided the computational power for executing the programs with 2 GB DDR3 dedicated memory.

The time-series data, after undergoing preprocessing, is suitably adapted for the LSTM model. An integral step in data preparation is standardization, as certain activation functions are sensitive to the scale of numerical values. Several standardization methods are explored, including MinMaxScaler, Quantile Transformer, and StandardScaler. After thoughtful evaluation, MinMaxScaler is determined to be the most appropriate approach for this particular study.

The LSTM model is designed with a stacked architecture featuring two LSTM layers, each consisting of 50 units. Subsequently, a dense layer with a single node is incorporated to generate the final output. The activation function employed in the LSTM layers is tanh, known for capturing complex patterns in sequential data. To optimize the model's performance during training, the Adam optimizer with a learning rate of 0.0004 is utilized, effectively controlling the weight updates and facilitating convergence.

For practical training, the mean absolute error (MAE) is selected as the loss function. The training process is carried out over 100 epochs, providing ample opportunity for the model to learn and adapt to the underlying patterns in the dataset. These parameters and hyperparameters are chosen by a systematic trial-and-error approach to achieve the most optimal results for the specific dataset and the forecasting task at hand. The number of epochs versus the loss function graph is a valuable reference to ascertain the appropriate number of epochs for training, preventing overfitting or underfitting of the model.

Throughout the implementation process, meticulous attention is devoted to refining the model architecture and fine-tuning the hyperparameters to enhance performance. Critical factors, such as the number of LSTM layers, the number of units in each layer, the learning rate, and the activation functions, are carefully evaluated to ensure the creation of a robust and effective LSTM model [60–65]. A similar process is applied to other models.

6.8. Error validation

The effectiveness of the proposed model is assessed using three key performance metrics: mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean square error (RMSE). These metrics are crucial in evaluating the disparities between predicted and actual values, providing valuable insights into the model's performance within the machine and deep learning realms.

Mean Absolute Error (MAE) is a fundamental measure to gauge the magnitude of differences between predicted and actual values. By calculating the average total errors across predictions and observations, MAE acts as the L1 loss function, making it particularly suitable for regression tasks. Its straightforward and quantitative nature enables researchers and practitioners to precisely evaluate the average magnitude of errors in regression and forecasting.

Mean Absolute Percentage Error (MAPE) is widely used in machine and deep learning to measure average absolute relative errors. One of its key advantages is its representation as a percentage, facilitating easy interpretation across diverse domains. Unlike metrics expressed in absolute terms, MAPE's relative nature removes the need for subject expertise, making it accessible to a broad audience, including non-technical stakeholders. As a result, MAPE offers a practical and intuitive means of assessing the model's performance.

Root Mean Square Error (RMSE) is considered one of the most dependable methods for assessing forecast accuracy. By quantifying the differences between estimated and actual measured values using the Euclidean distance, RMSE provides a comprehensive evaluation of the model's performance. Its sensitivity to outliers emphasizes the importance of accurate predictions, particularly in applications where precision is crucial. RMSE is extensively used in various fields, including educational monitoring and evaluating model performance throughout the training and validation.

Equations 12,13,14 represents the MAE, MAPE, and RMSE, respectively.

$$MAE = \frac{1}{n} \sum_{i=1}^{n} | A_i - F_i | \quad (12)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} | \frac{A_i - F_i}{A_i} | \quad (13)$$

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(A_i - F_i)^2}{n}} \quad (14)$$

Where,

A_i = Actual values

F_i = Forecasted values

n = Number of observations

7. Results and discussions

The prepared dataset trains and tests the models multiple times with different training and testing data ratios. The performance of the models is evaluated by measuring the errors in the forecasting task, specifically MAE, MAPE, and RMSE. To gain a comprehensive understanding of the model's performance and the efficacy of the dataset, variations in the errors across different temporal trends are considered. This includes analyzing the yearly and monthly variations of the errors, which are graphically represented in Figures 10 to 12 for yearly trends and Figures 13 to 15 for monthly trends. The yearly trend is examined over ten years (2023 to 2032), while the monthly trend focuses on 2024. These graphical representations offer insights into the fluctuations and patterns in the errors over time.

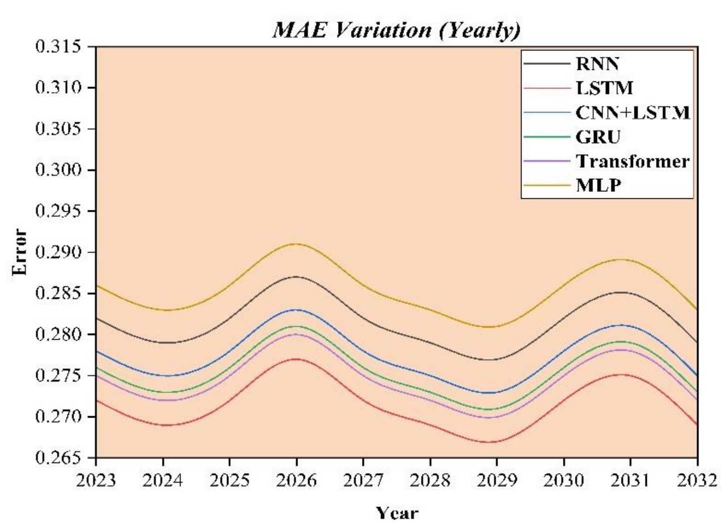

Fig. 10. MAE Variation (Yearly).

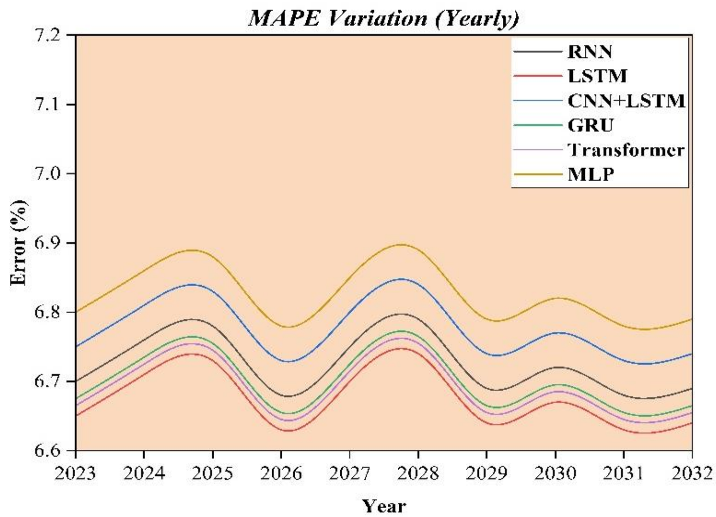

Fig. 11. MAPE Variation (Yearly).

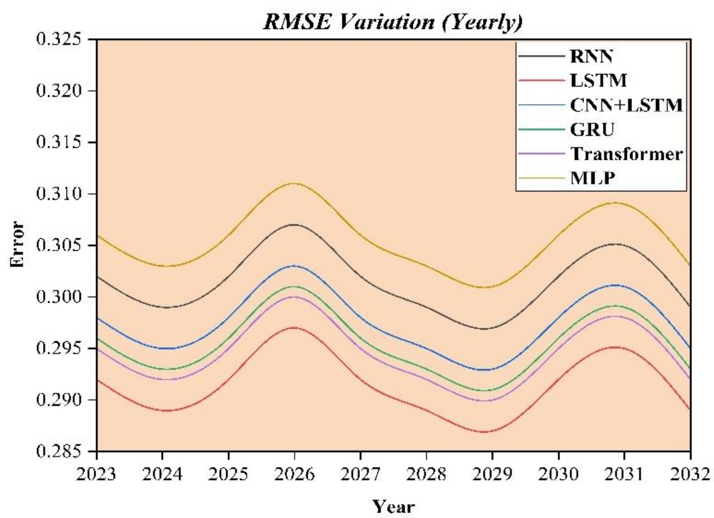

Fig. 12. RMSE Variation (Yearly).

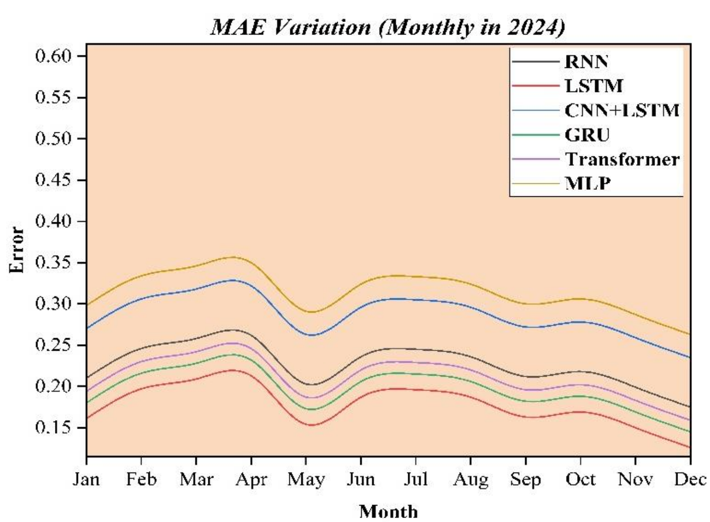

Fig. 13. MAE Variation (Monthly in 2024).

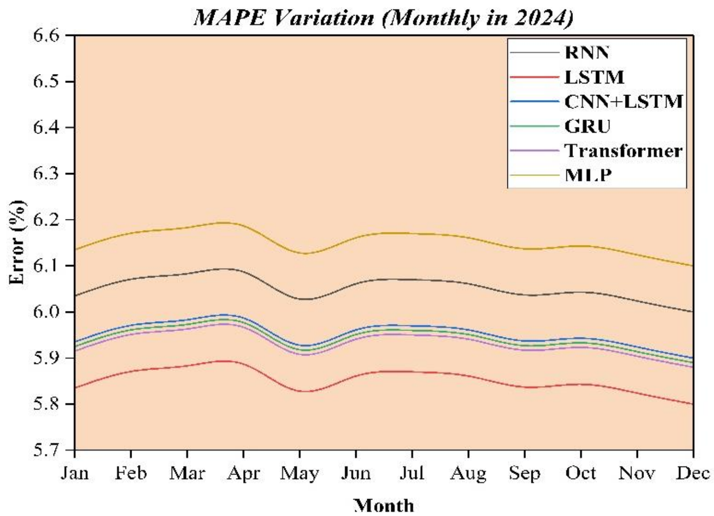

Fig. 14. MAPE Variation (Monthly in 2024)**.**

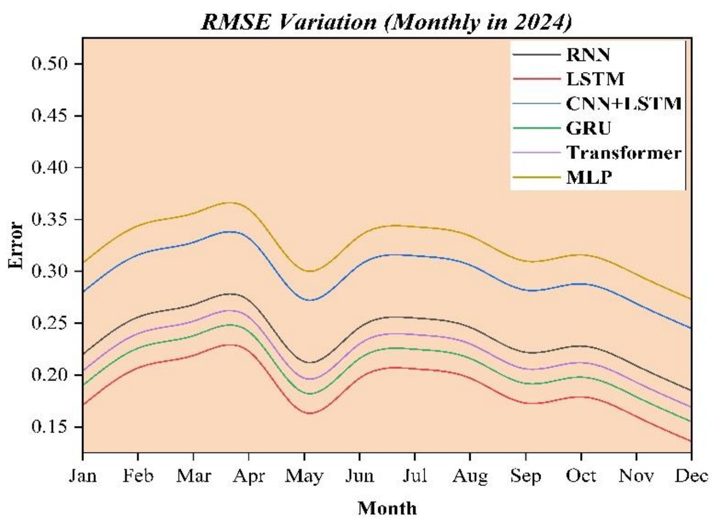

Fig. 15. RMSE Variation (Monthly in 2024)**.**

Based on the analysis, the average errors are calculated and are represented in Tables 1 and 2. Tables 1 and 2 illustrate the average errors of yearly and monthly forecasting, respectively.

Table 1
Average Error of Yearly Forecasting of Road Accidents.

Model	MAE (Mean Absolute Error)	MAPE (Mean Absolute Percentage Error)	RMSE (Root Mean Square Error)
RNN	0.281	6.724	0.301
LSTM	0.271	6674	0.291
CNN+LSTM	0.277	6.774	0.297
GRU	0.275	6.699	0.295
Transformer	0.274	6.689	0.294
MLP	0.285	6.824	0.305

Table 2
Average Error of Monthly Forecasting of Road Accidents.

Model	MAE (Mean Absolute Error)	MAPE (Mean Absolute Percentage Error)	RMSE (Root Mean Square Error)
RNN	0.225	6.050	0.235
LSTM	0.176	5.850	0186
CNN+LSTM	0.245	5.950	0.255
GRU	0.205	5.940	0.215
Transformer	0.195	5.930	0.205
MLP	0.265	6.150	0.275

The forecasting of road accidents is performed yearly and monthly, and the outcomes are visually presented in Figures 16 and 17, respectively. Data from ten years (2023 to 2032) is considered for the yearly trend, while the monthly trend focuses on the year 2024. These forecasting exercises aim to gain insights into the behavior of road accidents over time, considering cumulative effects influenced by various factors throughout the year. By analyzing the yearly variations, researchers can observe long-term patterns and trends in road accidents, identifying potential factors contributing to their fluctuations. On the other hand, monthly forecasting allows for a more granular examination of accident patterns within a year, enabling the detection of seasonality or specific temporal variations in accident occurrences. Through this comprehensive approach, a better understanding of road accident trends can be achieved, aiding in formulating effective road safety measures and policies.

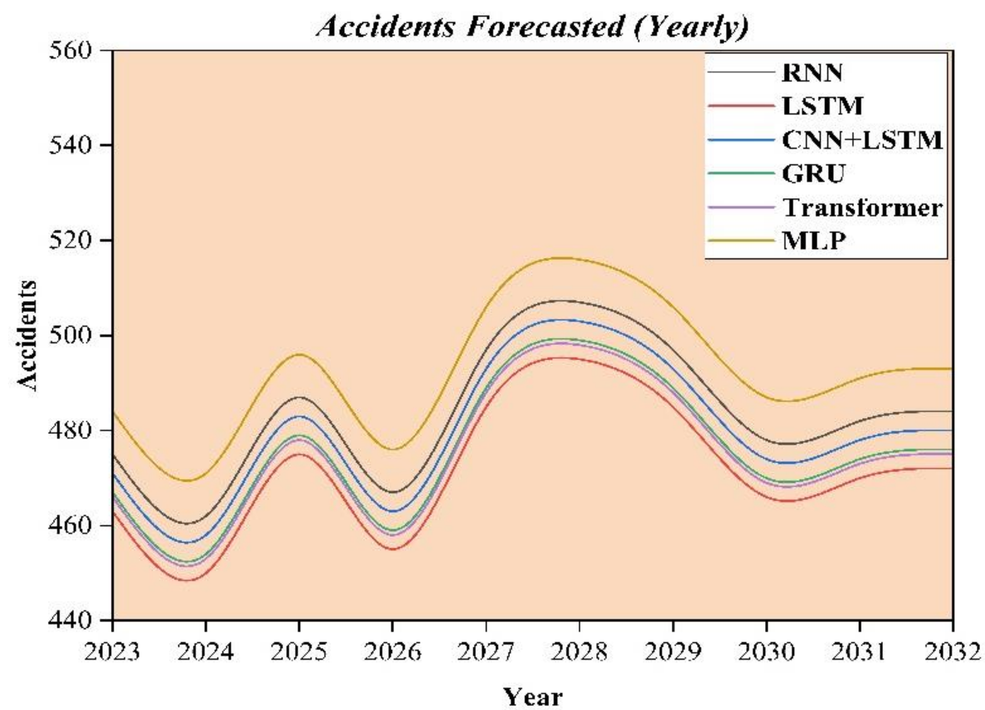

Fig. 16. Accident Forecasting (Yearly).

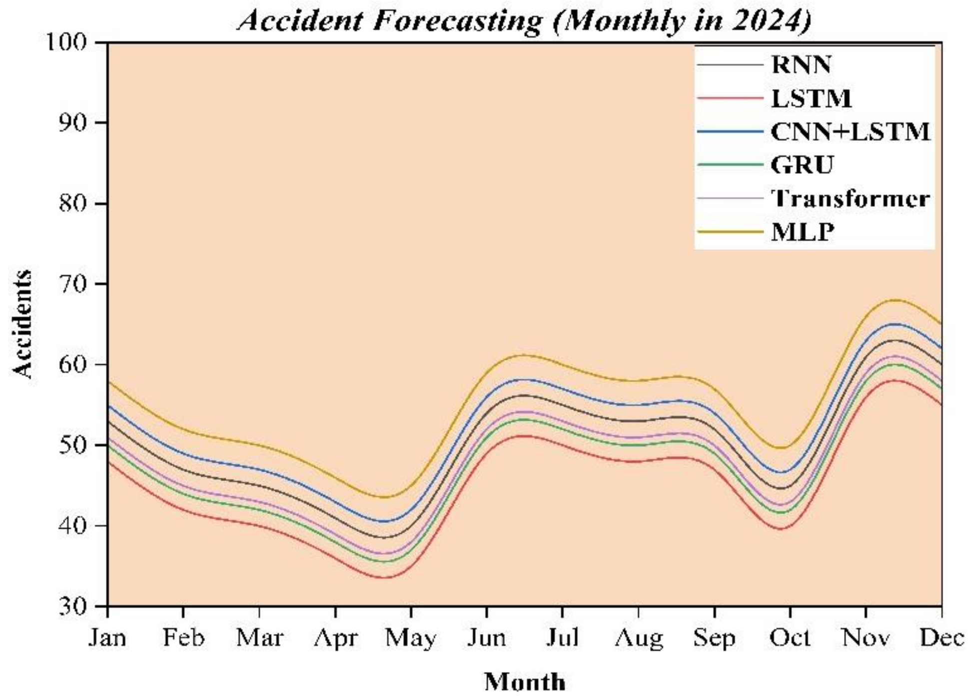

Fig. 17. Accident Forecasting (Monthly in 2024).

Examining temporal variation's influence on road accident forecasting reveals interesting insights, and the subsequent paragraphs delve into the results. The performance evaluation of deep learning models involves thoroughly analyzing the errors in the obtained results. Different types of errors are considered to compare the models' efficiency and determine the most suitable one for the study. As suggested by a standard guideline in the literature, the model with the lowest error is considered the most appropriate choice.

Upon analyzing Figures 10 to 15, a notable observation emerges: the variance of errors is smoother in monthly forecasting compared to yearly forecasting. This phenomenon is attributed to the nature of data variation. Monthly forecasting relies on more data points, contributing to a more refined and smoother error variation. By carefully examining Figures 10 to 15, along with Tables 1 and 2, a clear trend emerges: monthly forecasting exhibits fewer errors than yearly forecasting. Consequently, it can be confidently concluded that monthly forecasting yields superior results, indicating a better pattern recognition of accident occurrences. To gain further insights, Figure 16 showcases the forecasted accidents from 2023 to 2032, offering valuable information for devising strategies to improve the accident rate.

Figure 17 provides an intriguing observation: accidents tend to decrease from January to June as winter transitions into summer. During this period, enhanced visibility on major roads and improved driver performance contribute to declining accidents. However, an interesting contrast is observed in the urban regions of Hubballi-Dharwad, where visibility is compromised during winter due to fog development, leading to increased accidents. As the rainy season sets in from June to October, accidents surge. Hubballi-Dharwad's proximity to the Western Ghats attracts wind currents, resulting in heavy rainfall, impacting road quality, reducing visibility, and increasing accidents. Similar trends are observed from October to December, indicating the importance of weather-related considerations in accident forecasting.

Figures 10 to 15 and Tables 1 and 2 provide compelling evidence that the LSTM model outperforms other models in road accident forecasting based on error analysis. Its superior performance demonstrates the effectiveness and reliability of the LSTM model for this specific task.

Examining temporal variation's influence on road accident forecasting has brought forth valuable insights. The meticulous evaluation of deep learning models based on error analysis allows for identifying the most suitable model. Monthly forecasting exhibits smoother error variance and better results, enhancing pattern recognition of accident occurrences. The seasonal trends uncovered in Figure 17 underscore the significance of considering weather-related factors in accident forecasting. Ultimately, the LSTM model stands out as the top performer in road accident forecasting, validating its effectiveness and reinforcing its utility in this domain.

8. Conclusions

The primary objective of this research is to forecast road accidents by comparing different deep-learning models. The study has yielded significant observations that are crucial in concluding the findings.

Deep learning has emerged as a suitable approach, particularly in time series forecasting for road accidents. The inherent capability of deep learning models to handle large datasets has led to successful outcomes with minimal errors. This advantage has streamlined the training and testing processes, reducing the required time.

The research considered several deep learning models for time series forecasting of road accidents. Through error analysis, it was found that the LSTM model outperforms others, showcasing an RMSE of 0.291, MAE of 0.271, and MAPE of 6.674% in yearly variation. Furthermore, the LSTM model achieved an RMSE of 0.186, MAE of 0.176, and MAPE of 5.850% in monthly variation. These findings establish the effectiveness of LSTM in road accident forecasting.

Prior studies have emphasized using GRU and Transformer models to address the vanishing gradient problem, which has shown favorable results in various domains. However, the current error analysis demonstrates that LSTM, GRU, and Transformer models exhibit similar errors, with LSTM demonstrating the lowest. This suggests that the efficiency of the deep learning model is primarily determined by the characteristics of the chosen data for the study rather than the code structure.

The monthly forecasted results have unveiled a pattern where road accidents tend to increase from June to February, indicating a direct influence of climatic seasons on accident occurrences. To mitigate this, implementing measures such as reducing safe speeds on major roads, enhancing the visibility of red signals in adverse conditions, rerouting traffic, and increasing the duration of red signals can be beneficial in reducing accidents by minimizing traffic movement.

The forecasted accident data obtained from the deep learning models holds significant value for policymaking. These data-driven insights can be effectively utilized to conclude the study and further develop short-term and long-term policies to enhance safety and reduce the accident rate. With monthly forecasted data guiding short-term policies related to frequent road repairs, yearly forecasted data can inform long-term policies involving road expansion, construction of all-weather roads, planning of flyovers, implementing speed regulations in accident-prone areas, and more. The successful implementation of policies to improve the accident rate would contribute to overall city development, attracting new opportunities and investments.

It is important to note that the study focused on utilizing data from the urban regions of Hubballi-Dharwad for time series forecasting. However, incorporating data from the entire Dharwad district spanning 30 to 45 years would significantly enhance the forecasting comprehensiveness and granularity.

In conclusion, comparing deep learning models for road accident forecasting has resulted in valuable insights. The LSTM model has exhibited superior performance, and the research findings provide crucial information for policymakers to enhance safety, develop infrastructure, and improve the overall accident rate. To further enhance the accuracy and applicability of the forecasting models, it is recommended to expand the scope of the study by incorporating a larger dataset and extending the timeframe.

In summary, deep learning has proven to be a suitable approach for road accident forecasting, with LSTM demonstrating superior performance compared to other models. Policymakers can leverage the forecasted data to develop effective short-term and long-term policies to enhance safety and reduce accident rates. Additionally, the study underscores the potential benefits of incorporating a broader dataset and longer timeframe to improve forecasting accuracy and its practicality in policymaking. Ultimately, the insights gained from this research can contribute to city development, attracting new opportunities and investments.

Acknowledgments

The authors sincerely acknowledge the police officers and staff of the Hubballi Dharwad Police Commissioner's Office for their complete support in providing the necessary data and cooperation during the complete study.

Conflicts of interest

The authors declare no conflict of interest.

References

[1] Ministry Of Road Transport and Highways. New Delhi: 2021.

[2] Castillo-Manzano JI, Castro-Nuño M, López-Valpuesta L, Pedregal DJ. From legislation to compliance: The power of traffic law enforcement for the case study of Spain. Transp Policy 2019;75:1–9. https://doi.org/10.1016/j.tranpol.2018.12.009.

[3] Chand A, Jayesh S, Bhasi AB. Road traffic accidents: An overview of data sources, analysis techniques and contributing factors. Mater Today Proc 2021;47:5135–41. https://doi.org/10.1016/j.matpr.2021.05.415.

[4] Eskandari Torbaghan M, Sasidharan M, Reardon L, Muchanga-Hvelplund LCW. Understanding the potential of emerging digital technologies for improving road safety. Accid Anal Prev 2022;166:106543. https://doi.org/10.1016/j.aap.2021.106543.

[5] Sumit K, Ross V, Ruiter RAC, Polders E, Wets G, Brijs K. An Exploration of Characteristics and Time Series Forecast of Fatal Road Crashes in Manipal, India. Sustainability 2022;14:2851. https://doi.org/10.3390/su14052851.

[6] Cerqueira V, Torgo L, Soares C. Machine Learning vs Statistical Methods for Time Series Forecasting: Size Matters 2019.

[7] Athiappan K, Karthik C, Rajalaskshmi M, Subrata C, Dastjerdi HR, Liu Y, et al. Identifying Influencing Factors of Road Accidents in Emerging Road Accident Blackspots. Adv Civ Eng 2022;2022. https://doi.org/10.1155/2022/9474323.

[8] Pal R, Ghosh A, Kumar R, Galwankar S, Paul S, Pal S, et al. Public health crisis of road traffic accidents in India: Risk factor assessment and recommendations on prevention on the behalf of the Academy of Family Physicians of India. J Fam Med Prim Care 2019;8:775. https://doi.org/10.4103/jfmpc.jfmpc_214_18.

[9] Sangare M, Gupta S, Bouzefrane S, Banerjee S, Muhlethaler P. Exploring the forecasting approach for road accidents: Analytical measures with hybrid machine learning. Expert Syst Appl 2021;167:113855. https://doi.org/10.1016/j.eswa.2020.113855.

[10] Jha K, Sinha N, Arkatkar SS, Sarkar AK. A Comparative Study on Application of Time Series Analysis for Traffic Forecasting in India: Prospects and Limitations. Curr Sci 2016;110:373. https://doi.org/10.18520/cs/v110/i3/373-385.

[11] Mutangi K. Time series analysis of road traffic accidents in Zimbabwe 2015.

[12] Quddus MA. Time series count data models: An empirical application to traffic accidents. Accid Anal Prev 2008;40:1732–41. https://doi.org/10.1016/j.aap.2008.06.011.

[13] Commandeur JJF, Bijleveld FD, Bergel-Hayat R, Antoniou C, Yannis G, Papadimitriou E. On statistical inference in time series analysis of the evolution of road safety. Accid Anal Prev 2013;60:424–34. https://doi.org/10.1016/j.aap.2012.11.006.

[14] Halyal S, Mulangi RH, Harsha MM. Short-term passenger demand modelling using automatic fare collection data: a case study of Hubli-Dharwad BRTS. Adv Transp Stud 2023;59.

[15] Rabbani MBA, Musarat MA, Alaloul WS, Rabbani MS, Maqsoom A, Ayub S, et al. A Comparison Between Seasonal Autoregressive Integrated Moving Average (SARIMA) and Exponential Smoothing (ES) Based on Time Series Model for Forecasting Road Accidents. Arab J Sci Eng 2021;46:11113–38. https://doi.org/10.1007/s13369-021-05650-3.

[16] Caliendo C, Guida M, Parisi A. A crash-prediction model for multilane roads. Accid Anal Prev 2007;39:657–70. https://doi.org/10.1016/j.aap.2006.10.012.

[17] Kumar CN, Parida M, Jain SS. Poisson Family Regression Techniques for Prediction of Crash Counts Using Bayesian Inference. Procedia - Soc Behav Sci 2013;104:982–91. https://doi.org/10.1016/j.sbspro.2013.11.193.

[18] Karimnezhad A, Moradi F. Road accident data analysis using Bayesian networks. Transp Lett 2017;9:12–9. https://doi.org/10.1080/19427867.2015.1131960.

[19] Manaswi NK. RNN and LSTM. Deep Learn. with Appl. Using Python, Berkeley, CA: Apress; 2018, p. 115–26. https://doi.org/10.1007/978-1-4842-3516-4_9.

[20] Tokgöz A, Ünal G. A RNN based time series approach for forecasting turkish electricity load. 2018 26th Signal Process. Commun. Appl. Conf., IEEE; 2018, p. 1–4. https://doi.org/10.1109/SIU.2018.8404313.

[21] Shen Q, Wu Y, Jiang Y, Zeng W, LAU AKH, Vianova A, et al. Visual Interpretation of Recurrent Neural Network on Multi-dimensional Time-series Forecast. 2020 IEEE Pacific Vis. Symp., IEEE; 2020, p. 61–70. https://doi.org/10.1109/PacificVis48177.2020.2785.

[22] Sagheer A, Kotb M. Time series forecasting of petroleum production using deep LSTM recurrent networks. Neurocomputing 2019;323:203–13. https://doi.org/10.1016/j.neucom.2018.09.082.

[23] Chimmula VKR, Zhang L. Time series forecasting of COVID-19 transmission in Canada using LSTM networks. Chaos, Solitons & Fractals 2020;135:109864. https://doi.org/10.1016/j.chaos.2020.109864.

[24] Shin D-H, Chung K, Park RC. Prediction of Traffic Congestion Based on LSTM Through Correction of Missing Temporal and Spatial Data. IEEE Access 2020;8:150784–96. https://doi.org/10.1109/ACCESS.2020.3016469.

[25] Zheng J, Huang M. Traffic Flow Forecast Through Time Series Analysis Based on Deep Learning. IEEE Access 2020;8:82562–70. https://doi.org/10.1109/ACCESS.2020.2990738.

[26] Halyal S, Mulangi RH, Harsha MM. Forecasting public transit passenger demand: With neural networks using APC data. Case Stud Transp Policy 2022;10:965–75. https://doi.org/10.1016/j.cstp.2022.03.011.

[27] Qin D, Yu J, Zou G, Yong R, Zhao Q, Zhang B. A Novel Combined Prediction Scheme Based on CNN and LSTM for Urban PM 2.5 Concentration. IEEE Access 2019;7:20050–9. https://doi.org/10.1109/ACCESS.2019.2897028.

[28] Livieris IE, Pintelas E, Pintelas P. A CNN–LSTM model for gold price time-series forecasting. Neural Comput Appl 2020;32:17351–60. https://doi.org/10.1007/s00521-020-04867-x.

[29] Li P, Abdel-Aty M, Yuan J. Real-time crash risk prediction on arterials based on LSTM-CNN. Accid Anal Prev 2020;135:105371. https://doi.org/10.1016/j.aap.2019.105371.

[30] Liu Y, Wang Z, Yu X, Chen X, Sun M. Memory-based Transformer with shorter window and longer horizon for multivariate time series forecasting. Pattern Recognit Lett 2022;160:26–33. https://doi.org/10.1016/j.patrec.2022.05.010.

[31] Agga A, Abbou A, Labbadi M, Houm Y El, Ou Ali IH. CNN-LSTM: An efficient hybrid deep learning architecture for predicting short-term photovoltaic power production. Electr Power Syst Res 2022;208:107908. https://doi.org/10.1016/j.epsr.2022.107908.

[32] Mehtab S, Sen J. Analysis and Forecasting of Financial Time Series Using CNN and LSTM-Based Deep Learning Models, 2022, p. 405–23. https://doi.org/10.1007/978-981-16-4807-6_39.

[33] Zhou S, Guo S, Du B, Huang S, Guo J. A Hybrid Framework for Multivariate Time Series Forecasting of Daily Urban Water Demand Using Attention-Based Convolutional Neural Network and Long Short-Term Memory Network. Sustainability 2022;14:11086. https://doi.org/10.3390/su141711086.

[34] Li S, Yang H, Cheng R, Ge H. Hybrid deep learning models for short-term demand forecasting of online car-hailing considering multiple factors. Transp Lett 2024;16:218–33. https://doi.org/10.1080/19427867.2023.2175420.

[35] Wang D, Gan J, Mao J, Chen F, Yu L. Forecasting power demand in China with a CNN-LSTM model including multimodal information. Energy 2023;263:126012. https://doi.org/10.1016/j.energy.2022.126012.

[36] Yuan Z, Zhou X, Yang T. Hetero-ConvLSTM. Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., New York, NY, USA: ACM; 2018, p. 984–92. https://doi.org/10.1145/3219819.3219922.

[37] Zheng M, Li T, Zhu R, Chen J, Ma Z, Tang M, et al. Traffic Accident's Severity Prediction: A Deep-Learning Approach-Based CNN Network. IEEE Access 2019;7:39897–910. https://doi.org/10.1109/ACCESS.2019.2903319.

[38] Zhang Z, Yang W, Wushour S. Traffic Accident Prediction Based on LSTM-GBRT Model. J Control Sci Eng 2020;2020:1–10. https://doi.org/10.1155/2020/4206919.

[39] Yu L, Du B, Hu X, Sun L, Han L, Lv W. Deep spatio-temporal graph convolutional network for traffic accident prediction. Neurocomputing 2021;423:135–47. https://doi.org/10.1016/j.neucom.2020.09.043.

[40] Rajagukguk RA, Ramadhan RAA, Lee H-J. A Review on Deep Learning Models for Forecasting Time Series Data of Solar Irradiance and Photovoltaic Power. Energies 2020;13:6623. https://doi.org/10.3390/en13246623.

[41] Flores A, Tito H, Centty D. Comparison of Hybrid Recurrent Neural Networks for Univariate Time Series Forecasting, 2021, p. 375–87. https://doi.org/10.1007/978-3-030-55180-3_28.

[42] Zheng W, Chen G. An Accurate GRU-Based Power Time-Series Prediction Approach With Selective State Updating and Stochastic Optimization. IEEE Trans Cybern 2022;52:13902–14. https://doi.org/10.1109/TCYB.2021.3121312.

[43] Pirani M, Thakkar P, Jivrani P, Bohara MH, Garg D. A Comparative Analysis of ARIMA, GRU, LSTM and BiLSTM on Financial Time Series Forecasting. 2022 IEEE Int. Conf. Distrib. Comput. Electr. Circuits Electron., IEEE; 2022, p. 1–6. https://doi.org/10.1109/ICDCECE53908.2022.9793213.

[44] Shen L, Wang Y. TCCT: Tightly-coupled convolutional transformer on time series forecasting. Neurocomputing 2022;480:131–45. https://doi.org/10.1016/j.neucom.2022.01.039.

[45] Almuammar M, Fasli M. Deep Learning for Non-stationary Multivariate Time Series Forecasting. 2019 IEEE Int. Conf. Big Data (Big Data), IEEE; 2019, p. 2097–106. https://doi.org/10.1109/BigData47090.2019.9006192.

[46] Borghi PH, Zakordonets O, Teixeira JP. A COVID-19 time series forecasting model based on MLP ANN. Procedia Comput Sci 2021;181:940–7. https://doi.org/10.1016/j.procs.2021.01.250.

[47] Cicek E, Akin M, Uysal F, Topcu Aytas R. Comparison of traffic accident injury severity prediction models with explainable machine learning. Transp Lett 2023;15:1043–54. https://doi.org/10.1080/19427867.2023.2214758.

[48] Astarita V, Haghshenas SS, Guido G, Vitale A. Developing new hybrid grey wolf optimization-based artificial neural network for predicting road crash severity. Transp Eng 2023;12. https://doi.org/10.1016/j.treng.2023.100164.

[49] Jafarzadeh Ghoushchi S, Shaffiee Haghshenas S, Memarpour Ghiaci A, Guido G, Vitale A. Road safety assessment and risks prioritization using an integrated SWARA and MARCOS approach under spherical fuzzy environment. Neural Comput Appl 2023;35:4549–67. https://doi.org/10.1007/s00521-022-07929-4.

[50] Guido G, Haghshenas SS, Haghshenas SS, Vitale A, Gallelli V, Astarita V. Development of a Binary Classification Model to Assess Safety in Transportation Systems Using GMDH-Type Neural Network Algorithm. Sustainability 2020;12:6735. https://doi.org/10.3390/su12176735.

[51] Guido G, Haghshenas SS, Haghshenas SS, Vitale A, Astarita V, Haghshenas AS. Feasibility of Stochastic Models for Evaluation of Potential Factors for Safety: A Case Study in Southern Italy. Sustainability 2020;12:7541. https://doi.org/10.3390/su12187541.

[52] Guido G, Haghshenas SS, Haghshenas SS, Vitale A, Astarita V, Park Y, et al. Evaluation of Contributing Factors Affecting Number of Vehicles Involved in Crashes Using Machine Learning Techniques in Rural Roads of Cosenza, Italy. Saf 2022, Vol 8, Page 28 2022;8:28. https://doi.org/10.3390/SAFETY8020028.

[53] Guido G, Shaffiee Haghshenas S, Shaffiee Haghshenas S, Vitale A, Astarita V. Application of Feature Selection Approaches for Prioritizing and Evaluating the Potential Factors for Safety Management in Transportation Systems. Computers 2022;11. https://doi.org/10.3390/computers11100145.

[54] Guido G, Shaffiee Haghshenas S, Vitale A, Astarita V, Geem ZW. A Quantitative and Qualitative Review of the Role of Intelligent Transportation Systems in Road Safety Studies through Three Decades. J Soft Comput Civ Eng 2023;7:93–109. https://doi.org/10.22115/scce.2023.382320.1597.

[55] Census of India- Karnataka - Series 30 - Part XII A - District Census Handbook, Dharwad, Directorate of Census Operations, Karnataka. 2011.

[56] Pawluszek-Filipiak K, Borkowski A. On the Importance of Train–Test Split Ratio of Datasets in Automatic Landslide Detection by Supervised Classification. Remote Sens 2020;12:3054. https://doi.org/10.3390/rs12183054.

[57] Rácz A, Bajusz D, Héberger K. Effect of Dataset Size and Train/Test Split Ratios in QSAR/QSPR Multiclass Classification. Molecules 2021;26:1111. https://doi.org/10.3390/molecules26041111.

[58] Xu Y, Goodacre R. On Splitting Training and Validation Set: A Comparative Study of Cross-Validation, Bootstrap and Systematic Sampling for Estimating the Generalization Performance of Supervised Learning. J Anal Test 2018;2:249–62. https://doi.org/10.1007/s41664-018-0068-2.

[59] Dhruv P, Naskar S. Image Classification Using Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN): A Review, 2020, p. 367–81. https://doi.org/10.1007/978-981-15-1884-3_34.

[60] Li X, Zhang L, Li D, Guo D. Construction and Simulation of a Strategic HR Decision Model Based on Recurrent Neural Network. J Math 2022;2022:1–11. https://doi.org/10.1155/2022/5390176.

[61] Huang L, Cai T, Zhu Y, Zhu Y, Wang W, Sun K. LSTM-Based Forecasting for Urban Construction Waste Generation. Sustainability 2020;12:8555. https://doi.org/10.3390/su12208555.

[62] Hou J, Wang Y, Hou B, Zhou J, Tian Q. Spatial Simulation and Prediction of Air Temperature Based on CNN-LSTM. Appl Artif Intell 2023;37. https://doi.org/10.1080/08839514.2023.2166235.

[63] Meng X, Zhu T, Li C. Construction of perfect dispatch learning model based on adaptive GRU. Energy Reports 2022;8:668–77. https://doi.org/10.1016/j.egyr.2022.02.250.

[64] Reza S, Ferreira MC, Machado JJM, Tavares JMRS. A multi-head attention-based transformer model for traffic flow forecasting with a comparative analysis to recurrent neural networks. Expert Syst Appl 2022;202:117275. https://doi.org/10.1016/j.eswa.2022.117275.

[65] Liu X, Liang Y, Huang C, Hu H, Cao Y, Hooi B, et al. Do We Really Need Graph Neural Networks for Traffic Forecasting? 2023.