International conference on Advanced Techniques in Communication Networking and Automation

RESEARCH ARTICLE OPEN ACCESS

RAILWAY PASSENGER VOLUME PREDICTION USING LSTM NEURAL NETWORK

Vimal Kumar.S *, Venkatesh.S *, Sriram.B.R*, Yugendar.M.R* and **Mrs.K.Sathiya Bharathi,M.E,

*Student, Department of Electronics and Communication Engg, Anjalai Ammal Mahalingam Engineering College, Thiruvarur(Dt).
Email id: vimal282001@gmail.com

**Assistant Professor, Department of Electronics and Communication Engg, Anjalai Ammal Mahalingam Engineering College, Thiruvaru(Dt).
Email id: sathiya2188@gmail.com

ABSTRACT

The growth of the railway passenger transport sector depends on accurate predictions of passenger volume. A neural network-based LSTM prediction model is created using railway data, using national railway volume of passengers as an example. The goal of the current project is to forecast future train passenger volume using the LSTM (Long Short-Term Memory) neural network. In order to find connections and trends in the data that can aid in forecasting future passenger traffic, the LSTM algorithm is trained using previous railway passenger data. The model makes use of sophisticated sequence modelling approaches to identify persistent dependencies in the information at hand. The study's findings show how well the model developed by LSTM predicts the number of passengers travelling by train.

Keyword: LSTM neural network

I. INTRODUCTION

In the current day, the railway plays an ever-more important role in the development of economies and has a significant influence on societal advancement. The railway has many advantages over other forms of transportation, including long travel distances, powerful transport capacities, resistance to climatic restrictions, high levels of safety, and inexpensive costs, among others. Predicting rail passenger volume is useful for determining the future course of rail passenger volume and is essential to the growth of the rail passenger transport sector. For railway operators, estimating the number of people who will travel upon a specific day or at a particular hour is a major task. It incorporates forecasting, which enables business owners to maximise resources, increase consumer satisfaction, and raise service quality as a whole. Estimating the number of passengers on trains is a critical task for railways operators to complete in order to meet customer demand while optimising resource use. The intricate structure of the demand-influencing factors, such as timing, fluctuations in demand, weather, as well as other extrinsic factors, makes forecasting passenger volume challenging. Conventional statistical approaches have been used to anticipate passenger volume, but they frequently have trouble capturing the complex nonlinear interactions between the different demand-influencing components.As a result, there is growing interest in using machine learning techniques, such as neural networks, to forecast passenger volumes. The prediction of passenger volume using LSTM, or long-short-term memory, neural networks has shown promise in addressing time-series forecasting challenges. Recurrent neural networks of the LSTM type can be helpful for forecasting future passenger volume utilising historical data because they may recognise dependencies that persist in time series information. The flexibility of LSTM networks to incorporate many inputs, including historical visitor volume data and other demand-influencing variables like the climate, holidays, and special events, is one of the key benefits of employing them to anticipate passenger volume.. These inputs enable LSTM networks to produce predictions that are more accurate than those made by conventional statistical models. However, there are drawbacks to utilising LSTM networks to anticipate passenger volume. One difficulty is that a sizable amount more historical data is required in order to develop the model successfully. The need to carefully choose and preprocess the model's inputs in order to ensure they accurately capture the pertinent aspects that drive demand presents an additional challenge. Overall, LSTM neural networks show promise for improving the accurateness of passenger volume forecast in the railway business, despite the difficulties. Railway operators may more effectively manage their resources, enhance their services, and enhance the overall experience of travellers by precisely projecting the flow of passengers.

II. LITERATURE SURVEY

Given its crucial significance in enhancing the effectiveness of the rail transport system, railway passenger traffic prediction has become a prominent area of study in recent years. Numerous theories have been put up in the literature to predict the number of passengers using trains, including statistical methodologies, machine learning strategies, and deep models for learning. However, there are still a number of pressing issues with regard to an expected volume of train passengers. The first issue is related to the complicated and shifting character of demand for passengers, which is affected by a variety of variables including weather, social gatherings, holidays, and economic conditions. These elements make it impossible to anticipate passenger volume accurately.. The second difficulty is the presence of outliers or noise in the information, which can significantly reduce the prediction's accuracy. By using the Particles Swarm Optimisation (PSO) algorithm to optimise the hyperparameters linked to the Generalised Regression Neural Networks (GRNN) model, a PSO-GRNN model for railways freight volume prediction aims to increase the forecasting's accuracy.Complexity overfitting, hyperparameter reliance, lack of accountability, limited data, and limited relevance are drawbacks for this paper.Due to their capacity for identifying and recollecting ongoing relationships in the data, lstm will emerge. Passenger Flow Forecasting of Combined Passenger Terminal was another study.

Based on K-Means-GRNN, the study groups travellers into groups according to their travel preferences and then uses the GRNN algorithm to forecast Passengers . Challenges, however, include interpretability issues, potential clustering errors, and real-time updating limitations.With the capacity to handle multiple group membership, real-time adjustments, and other issues, lstm has found solutions. Predictions are more dependable and consistent when they are robust to beginning conditions.in a separate study, rail traffic flow forecasting using ARIMA A model is used to examine and foretell the characteristics of a station's passenger flow. But it does have some restrictions, such as the linearity, stationarity, and normalcy of error assumptions, which might not always hold true in practical applications.In another study, a GBDT-enabling algorithm is used to forecast the number of passengers travelling by train between Beijing and Shanghai. This model might be harder to interpret than more straightforward models like regression, logistic regression, or decision tree models, which could be detrimental to its value.

III. LSTM NEURAL NETWORK

Recurrent neural networks of the the LSTM algorithm (Long Short-term Memory) variety are designed to recognise dependence over time in sequential input. In 1997, Hochreiter and Schmidhuber came up with the idea. Utilising a memory cell, LSTM can choose what data to update or delete as needed while yet preserving information over time. The first figure depicts the LSTM Neural Network's structural layout.The input gate, forget gate, and output gate are the three gates that control the cell. These gates acquire the knowledge of which data to accept and refuse as it enters the cell. LSTM cells function using a number of different parts. The input gate chooses the most important information given the input and gives each input element a value that ranges from 0 to 1, with 1 signifying relevance and 0 signifying insignificance. By using the tanh activation function, the candidate value adds to the cell's memory and compress the newly discovered data to be stored there to a value among -1 and 1. The forget gate determines which data should be removed from a memory cell by giving each component of the storage cell a value ranging from 0 to 1, with 1 signifying retention and 0 signifying removal. In order to update the cell state, the outputs of the input gates gates, forget gate, and candidate value are combined. The previous cell state is multiplied by the output of the forget gate, and the input gate input is divided by the potential value and added to it. Finally, based on the state of the cell at that moment, the output gate chooses what output to produce. Each component of a memory cell is given a value from 0 to 1, with 1 denoting importance and 0 denoting insignificance. The cell state is multiplied by the output gate output, and an activation function such as tanh or sigmoid is employed.

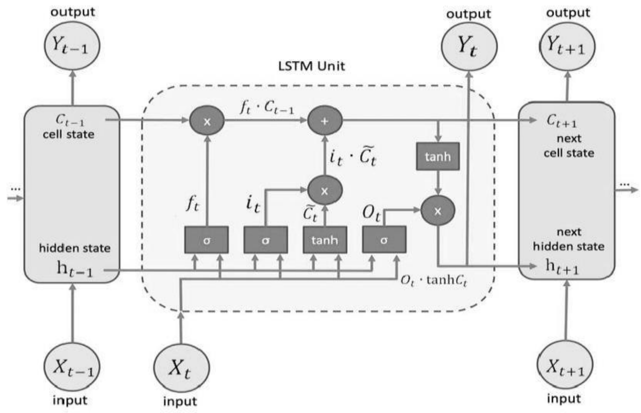

Figure1.Structure of LSTM Neural Network

The sigmoid & tanh functions are used in a LSTM neural network in order to control input flow and modify cell state. The function known as the sigmoid determines whether the data is retained or abandoned. It works by interpreting the input value as a stochastic value and compressing it into a range between 0 and 1. The way it works can be described as a gating system that allows only important data to flow through while blocking unnecessary data. The sigmoid function has what follows :

σ(x) = 1 / (1 + e^(-x))

The "forget gate" of an LSTM uses the sigmoid function to determine which information from the cell state to ignore.A previously unknown piece of knowledge that can be included in the cell state is represented by the candidate state, which is created using the tanh function. The tanh function works by condensing the value of the input into a range that is limited to between -1 and 1. As it can efficiently capture non-linear characteristics in the input data and provide a new state to be incorporated with the present cell state, this function is an invaluable asset to the LSTM. The tanh function has the following formula:

tanh(x) = (e^x - e^(-x)) / (e^x + e^(-x))

The tanh function performs two functions: it is an essential part of the "input gate" that creates the candidate state and is an essential part of the "output gate" that determines what information should be sent from the cell state. The cell state is continuously updated through these unique gates and the related functions, carefully safeguarding important data while eliminating unimportant data. The complicated the LSTM algorithm mechanism enables it to effectively handle persistent dependencies contained in sequential data, improving performance on a variety of challenging tasks.

IV. PREDICTION METHODOLOGIES

A. DATA PROCUREMENT

The task of gathering statistics about the number of passengers travelling on a specific day, time of day, destination, and season is required for the endeavour of accumulating data for railway patronage projection using LSTM (Long Short-Term Memory) neural network. Every record is recorded across a period of time as time-series data, which is how the collected data is frequently presented.Data scaling, categorical variable encoding, and splitting the data into sets for testing and training are all possible components of this preprocessing step. The LSTM model is then created, which uses the historical data on passenger volume as input and predicts future passenger volume. The training data is used to train the model, and the testing data is used to evaluate it. The goal is to create a model that predicts future passenger volumes with accuracy.

B. DATA STANDARDIZATION

In order to ensure that the collected data is suitable for usage with the LSTM model, data preparation is a crucial component. Data Cleaning, Data Scaling, Data Transformation, Dataset Splitting, and Data Enhancement are the crucial processes in the data preprocessing phase. The accuracy and durability of LSTM-based railway passenger volume forecast depend greatly on data preparation as a whole. The goal is to transform the raw data into an arrangement that the model developed by LSTM can use to efficiently provide accurate forecasts.

C. LSTM MODEL CONSTRUCTION

LSTM Model Development: Create an LSTM model using a state-of-the-art deep learning package, such as TensorFlow or Keras. The previous passenger volume data would be the model's input, and the estimated visitor volume for a later time would be the model's output

D. MODEL TRAINING

Model training is the procedure of introducing the training dataset to the artificial intelligence algorithm to discover the complex patterns hidden in the data. When using LSTM models, the model's algorithm is trained using historical passenger volume data in order to predict future visitor volume data.

E. MODEL APPRAISAL

Examining the effectiveness of the model that was trained utilising the testing dataset is the process of model appraisal. Various measures are used to assess the model's accuracy, including Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). These measurements measure the discrepancy between actual and projected values.

F. MODEL DEPLOYMENT

The practise of using a trained and verified model to predict new data is known as "model deployment." The model is shared and integrated into a larger structure, such as a web-based programme or mobile app, after it has been verified and evaluated in order to forecast fresh data in real-time.

V. EXPERIMENTAL ANALYSIS

A. EXPERIMENTAL ARENA

The LSTM neural network was implemented in this study using Keras as the experimentation platform. The use of Keras significantly reduces the coding burden connected with neural network techniques. Additionally, Keras' ease of use makes it a flexible tool for a wide range of applications. The main benefit of Keras is its ability to condense complex Tensorflow syntax into manageable lines of code. The a tool called Min module from the sklearn package was loaded, and the fit_transform() method was called to standardise the training data, the MinMaxScaler module from the sklearn library was imported, and the fit_transform() function was invoked.

1. REQUISITE SOFTWARE
   - Operating System : 64-bit Microsoft Windows 11, 10, 8, 7 (SP1)
   - Frond End : Python
   - Tool : Python 3.7
   - IDE : Visual Studio Code

B. CRITERION FUNCTION AND VARIATIONAL PARAMETEROPTIMIZER

The error function and optimisation procedure are used by neural networks constructed with LSTM for rail passenger volume prediction to assess and enhance the forecasting accuracy of the model. The most popular error operates for regression problems, such as predicting passenger volume, is the Mean Squared Error (MSE), which is expressed as: MSE = 1/N * (y - y_hat)2, where MSE stands for the Mean Squared Error, N stands for the overall number of simulated instances, y stands for the actual passengers volume, and y_hat stands for the anticipated passenger volume.

To reduce the MSE error function, the LSTM network's biases and weights are modified during the optimisation phase. In LSTM networks, the Adam optimizer is an algorithm which modifies the weights as well as the biases determined by both the initial and second moments that define the gradients

for regression problems. Specifically, the optimizer updates the weights and biases according to the following equations:

m = β1 * m + (1-β1) * dw

v = β2 * v + (1-β2) * dw^2

w = w - α * m / (sqrt(v) + ε)

where dw is the advanced gradient of the power source error operate with respect to the weights, m and v indicate the first and subsequent moments of the gradients, respectively, 1 and 2 denote the one that powers the decay rates for these moments, is the learning rate, is a small constant to prevent division by nothing, and w symbolises the weights of the network.

The LSTM network's algorithm can be trained on historical data on passenger volume along with other relevant factors like day of the week, hour of day. They weather conditions, etc holidays to produce accurate forecasts for upcoming passenger volumes.For forecasting the number of passengers travelling by train, LSTM neural networks frequently use the MSE error rate and Adam optimizer. During the optimization process, the network's weights and biases are adjusted to minimize the MSE error function, while the network can be trained on historical passenger volume data and other pertinent variables to generate accurate predictions for future passenger volumes.

C. ACTIVATION FUNCTION

An LSTM neural network cannot function properly without activation functions. To calculate a nerve cell's level of engagement or firing rate, a mathematical operation known as the activation function is applied to the nerve cell's output in the network. The network gains non-linearity through functions for activation, which enables it to describe complex connections among inputs and outputs.

The output of an LSTM cell in a network using LSTM is frequently subjected to the activation function before being passed on to the subsequent cell or its results layer. The hyperbolic tangential (tanh) or sigmoid functions are frequently used as activation functions in LSTM cells. Through controlling how frequently the cell's memory state is updated, these functions assist in modulating the flow of knowledge through the cell.

The input is mapped by the tanh function to a range among -1 and 1, while the input is mapped by the sigmoid function to a range of zero to one.

The exact task at hand and the type of data being processed determine which activation function should be used. Overall, the activation processes are essential for an LSTM neural network's smooth operation and are crucial for making accurate predictions.

D. ESTABLISHMENT OF LSTM NEURAL NETWORK

The creation of an LSTM neural network requires the collection and pre-processing of data, the formulation of the network's design, the conditioning of the network on the data, and the evaluation of the network's effectiveness. A loss function and optimisation algorithm condition the network, and its structure is adjusted to strike an appropriate equilibrium between complexity and accuracy. The network can be deployed for quick prognostication or other tasks after it has been trained and certified.

E. EXPERIMENTAL RESULTS AND ANALYSIS

In this study, it is crucial to preprocess the raw data before training the neural network using the LSTM on data from railway passenger flow. In order to prevent the input data from being excessively large or too tiny, which would lead to a badly trained model, the LSTM neural network's algorithm often requires normalised data as input. The range of normalised data is -1 to 1. Import the a min-max module using the sklearn toolbox, then use the fit_transform() method to normalise the training data.Let's assume that the input statistics for rail passenger volume are displayed in Figure 2.

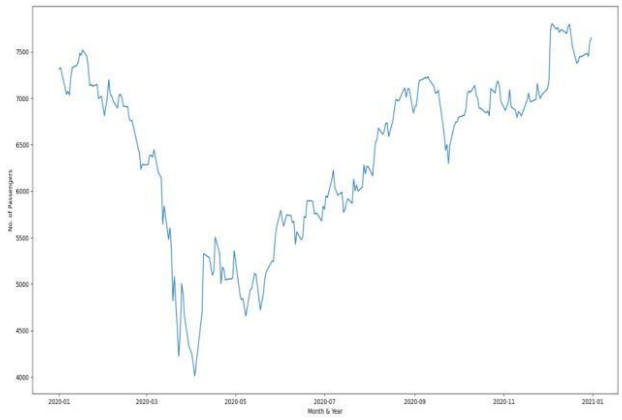

Fig 2,input statistics for rail passenger volume

When engaging in machine learning training, we must reframe the problem as a supervised learning problem. Generally, neural network data can be categorized into two groups: input (X) and output (Y). The neural network must discover the relationship between the observation at the previous time step (t-1) and the current time step (t). Using the shift function of the Pandas library in Python can facilitate the conversion process. according to research , outcome for railway passenger volume prediction in LSTM Neural Network with parameters are shown in figure.2

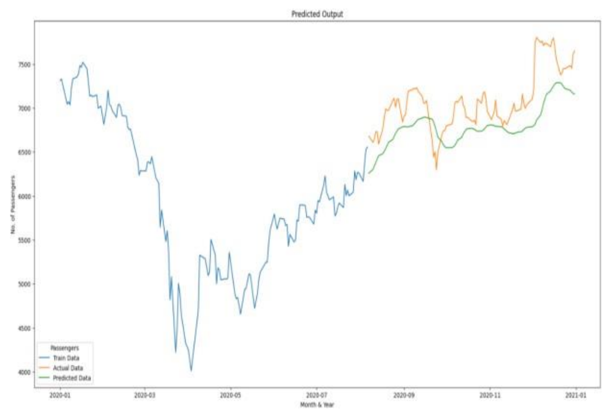

Figure 3. Prediction outcome of LSTM neural network

VI. CONCLUSION

Conclusively, employing LSTM neural network for railway passenger volume prediction can amplify the efficacy and efficacy of railway administration. By leveraging the historical passenger data, the model can prognosticate future passenger volumes with great precision, enabling railway authorities to judiciously strategize and allocate resources.The LSTM algorithm has a unique architecture with memory cells that preserve the past data and a set of gates that regulate the information flow, enabling the model to detect temporal patterns and trends in the passenger data, thereby generating accurate prognostications.It's vital to note that unforeseen events such as natural calamities or unforeseeable changes in travel patterns can influence the accuracy of the predictions. Therefore, utilizing the LSTM predictions as a tool to aid in decision-making and resource allocation can help railway authorities to stay abreast and proactive in managing passenger volumes.

VII. ACKNOWLEDGEMENT

We would like to extend our sincere gratitude to everyone who helped the study on railway passenger traffic forecasting using neural networks based on LSTM come to a successful conclusion.

First and foremost, we would like to express our sincere gratitude to our Project Guide for providing unwavering support, wise direction, and ongoing inspiration throughout the project. We are incredibly grateful for her vast knowledge and astute ideas, which enabled us to complete this assignment successfully.

We owe a debt of gratitude to our teammates for their continuous collaboration, support, and priceless input that improved our work.

REFERENCES

1. He, Z., Jiang, J., Cao, B., Zhang, X., & Xu, C. (2020). Prediction of Metro Passenger Volume Based on LSTM Neural Network Model. Mathematical Problems in Engineering, 2020, 1-12.

2. Yu Wang , Zhan Lin, Liang Wang , Hongye Wang , Junfeng Zhang .(2020) Prediction of Railway Passenger Volume. IEEE 9th Joint International Information Technology and Artificial Intelligence Conference (ITAIC).
3. Zeng, L., Liu, Y., & Gao, Z. (2020). A Novel Model for Urban Rail Transit Passenger Flow Prediction Using Deep Learning with Spatial-Temporal Attention Mechanism. Transportation Research Part C: Emerging Technologies, 111, 364-382.
4. Han, S., Luo, W., & Liu, X. (2019). Forecasting subway passenger flow using a long short-term memory neural network with a residual network. IEEE Transactions on Intelligent Transportation Systems, 20(5), 1656-1667.
5. Chen, L., Ma, Y., Xu, J., & Xu, L. (2019). Deep learning-based method for predicting urban rail transit passenger flow with the combination of multiple data sources. Transportation Research Part C: Emerging Technologies, 104, 235-252.
6. Luo, W., Han, S., & Guo, S. (2019). Prediction of subway passenger flow with a novel LSTM model. IEEE Transactions on Intelligent Transportation Systems, 20(7), 2564-2574.

7. Huang, S., Zhang, Y., & Wang, Y. (2019). Short-term passenger flow prediction using an improved LSTM model with a combination of external factors. Journal of Advanced Transportation, 2019, 1-15.
8. Liu, X., Luo, W., & Han, S. (2019). A deep learning approach to short-term subway passenger flow forecasting. IEEE Transactions on Intelligent Transportation Systems, 21(4), 1444-1454.
9. Jiang, J., He, Z., Cao, B., Zhang, X., & Xu, C. (2019). Short-term metro passenger volume prediction based on LSTM with multiple data sources. Journal of Advanced Transportation, 2019, 1-12.
10. Xia, Y., Huang, D., & Nie, C. (2018). Short-term passenger flow prediction for urban rail transit using a novel LSTM- based hybrid model. Journal of Advanced Transportation, 2018, 1-13.
11. Lai, K. K., Ng, A. W., Chan, F. T., & Lam, H. Y. (2018). Hybrid LSTM-DPSO for short-term metro passenger flow prediction. Transportation Research Part C: Emerging Technologies, 91, 224-238.
12. Zhao, J., Liu, Q., & Ding, Y. (2018). A hybrid model combining LSTM network with ARIMA for air passenger volume forecasting. Journal of Advanced Transportation, 2018,
13. Guo, S., Wu, C., Huang, J., & Luo, W. (2018). Short-term passenger flow prediction for metro systems using deep LSTM recurrent networks. Transportation Research Part C: Emerging Technologies, 94, 276-291.
14. Zhou, X., Wang, J., Chen, L., & Chen, M. (2018). A multi-source fusion method for short-term traffic flow prediction based on LSTM. IEEE Transactions on Intelligent Transportation Systems, 19(8), 2473-2482.
15. Yao, Q., Guo, S., & Luo, W. (2018). A hybrid model of LSTM and ARIMA for short-term metro passenger flow prediction. IEEE Access, 6, 53619-53627.
16. Zhang, Y., Huang, S., & Wang, Y. (2018). Short-term metro passenger flow prediction using an improved LSTM model. Journal of Advanced Transportation, 2018, 1-14.
17. Zhang, S., Cheng, Y., & Ren, Y. (2018). A deep learning based model for short-term urban rail transit passenger flow prediction. Transportation Research Part C: Emerging Technologies, 94, 333-349.
18. Wu, J., Yang, D., Wu, J., & Ye, M. (2018). Urban rail transit passenger flow prediction using a novel deep learning approach. Transportation Research Part C: Emerging Technologies, 97, 219-236.
19. Liu, Y., Li, Q., Li, Y., & He, Y. (2018). Short-term metro passenger flow prediction using an improved LSTM model with external information. Journal of Advanced Transportation, 2018, 1-12.
20. Wang, D., Guo, J., Liu, J., & Cai, J. (2017). Deep learning for short-term passenger flow prediction in transportation hub. Neurocomputing, 267, 378-385.
21. Chen, T., & Liao, Q. (2017). Short-term metro passenger flow prediction with deep learning methods. Journal of Advanced Transportation, 2017