International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 12 | December - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

Traffic Flow Forecast using Time Series Analysis based on Machine Learning Algorithm

KARTHIKEY JAYACHANDRA. V
B. Tech
School of Engineering
Computer Science-(AI&ML)
Malla Reddy University, India
2111CS020203@mallareddyuniversity.ac.in

KARTHIK YADAV. T
B. Tech
School of Engineering
Computer Science-(AI&ML)
Malla Reddy University, India
2111CS020201@mallareddyuniversity.ac.in

KARTHIKEYA. M
B. Tech
School of Engineering
Computer Science-(AI&ML)
Malla Reddy University, India
2111CS020202@mallareddyuniversity.ac.in

SHAMVITHA. K
B. Tech
School of Engineering
Computer Science-(AI&ML)
Malla Reddy University, India
2111CS020204@mallareddyuniversity.ac.in

KARTHIK REDDY. CH
B. Tech
School of Engineering
Computer Science-(AI&ML)
Malla Reddy University, India
2111CS020199@mallareddyuniversity.ac.in

KARTHIK REDDY. J
B. Tech
School of Engineering
Computer Science-(AI&ML)
Malla Reddy University, India
2111CS020200@mallareddyuniversity.ac.in

Guide: Prof.Divya Bharathi
Assistant Professor
School of Enigineering
Computer Science (AI&ML)
Malla Reddy University, India

ABSTRACT--- Time series forecasting plays a critical role in predicting future trends and making informed decisions across various domains. In this research paper, we present a multi-model approach to improve the accuracy and robustness of time series forecasting. Our methodology involves integrating two powerful models, Long Short-Term Memory (LSTM) and Seasonal Autoregressive Integrated Moving Average (SARIMA), to capture complex temporal patterns and seasonality in the data.

We leverage real-world traffic data from different junctions and employ a comprehensive pipeline for data preprocessing, including data selection, resampling, and sequence generation. Our study focuses on four junctions, and for each junction, we individually train LSTM models, followed by SARIMA models for comparison.

The LSTM models are designed to capture long-term dependencies in the time series data, while the SARIMA models are tailored to handle seasonality and autocorrelation. We implement a custom sequence generation process and split the dataset into training and testing sets for model evaluation. The LSTM models are trained using Pytorch, optimizing them for accurate short-term predictions.

Our results show that the combination of LSTM and SARIMA models yields superior forecasting performance compared to using each model individually. We present a detailed analysis of the forecasted results, including Root Mean Square Error (RMSE) calculations and visualizations to demonstrate the effectiveness of our multi-model approach.

This research contributes to the field of time series forecasting by showcasing the benefits of combining deep learning and classical statistical methods. The proposed approach provides a flexible and robust framework applicable to various time series prediction tasks, offering improved accuracy and reliability in forecasting future trends.

I. INTRODUCTION

The introduction delves into the pivotal role of intelligent transportation systems in urban planning, with a particular focus on the necessity of accurate traffic flow prediction for optimizing various aspects of traffic management. The objectives of the study are outlined, centering on the evaluation of LSTM and SARIMA models for traffic flow prediction. The comparative analysis extends to different junctions, aiming to gauge the models' accuracy variations. Additionally, the study seeks to unravel the strengths and limitations inherent in each model, specifically in capturing temporal patterns.

The motivation behind this research lies in the broader goal of enhancing traffic signal timings, congestion management, and overall transportation efficiency. By scrutinizing the performance of advanced prediction models, the study aims to provide valuable insights into improving the efficacy of urban traffic management systems. The outcomes are expected to contribute significantly to2 addressing practical challenges in traffic forecasting, paving the way for more robust and reliable methods. Ultimately, the research endeavors to facilitate advancements in urban transportation, fostering smarter and more responsive systems that adapt to dynamic traffic conditions. Through this exploration, the study aspires to play a vital role in the ongoing evolution of intelligent transportation systems, offering tangible solutions for more effective urban mobility.

II. LITERATURE SURVEY

Intelligent Transportation Systems (ITS) are integral to urban planning, with traffic flow prediction standing as a crucial component. The literature extensively explores forecasting methodologies, with a notable focus on Long Short-Term Memory (LSTM) and Seasonal Autoregressive Integrated Moving Average (SARIMA) models.

LSTM in Traffic Flow Prediction: LSTM, a type of recurrent neural network, has gained popularity for its ability to capture long-term dependencies in sequential data. In traffic flow prediction, LSTM demonstrates efficacy in learning intricate patterns from historical traffic data, allowing for accurate predictions.

SARIMA in Traffic Flow Prediction: Seasonal Autoregressive Integrated Moving Average (SARIMA) models, rooted in time-series analysis, address the temporal aspects of traffic patterns. Particularly adept at handling seasonal variations, SARIMA contributes to robust predictions in scenarios where temporal dependencies play a crucial role.

Comparative Studies: Numerous studies have compared the performance of LSTM and SARIMA models in traffic flow prediction. While LSTM excels in capturing complex patterns and non-linear relationships, SARIMA's strength lies in handling seasonality and short-term variations. Comparative analyses provide valuable insights into the suitability of each model under varying conditions.

Challenges and Limitations: Despite their success, challenges persist in deploying these models across diverse urban landscapes. The variability in traffic patterns, unexpected events, and model interpretability pose ongoing challenges. Understanding and addressing these limitations are imperative for ensuring the practical applicability of the models.

Research Gaps: While the literature offers substantial insights, notable research gaps persist. Specific applications of LSTM and SARIMA models to diverse urban contexts demand further exploration. Additionally, a comprehensive analysis of the strengths and limitations of each model, especially in capturing temporal patterns, remains an area for refinement.

Conclusion: In conclusion, the literature review establishes the significance of LSTM and SARIMA models in traffic flow prediction. However, it underscores the need for continued research to address challenges and refine these models for real-world, dynamic traffic scenarios. The subsequent empirical evaluations in this study aim to contribute substantively to this evolving field of intelligent transportation systems.

III. PROBLEM STATEMENT

Urban traffic management is a critical facet of modern city planning, demanding accurate and efficient predictive tools. The challenge lies in optimizing traffic flow, signal timings, and congestion management to enhance overall transportation efficiency. The existing urban infrastructure is intricate, with dynamic and often unpredictable traffic patterns. Traditional methods are proving insufficient to cope with the complexity, necessitating advanced predictive models.

The problem at hand is the need for a robust and adaptable traffic flow prediction system. Accurate forecasts are imperative for effective urban planning, enabling timely adjustments in traffic signal timings and congestion mitigation. The inadequacy of conventional models prompts the exploration of cutting-edge methodologies.

The research seeks to contribute to the development of intelligent transportation systems, offering insights into the most effective models for accurate and adaptable traffic flow predictions in diverse urban landscapes.

IV. METHODOLOGY

Traffic flow prediction relies on robust methodologies to extract meaningful patterns from complex datasets. This section outlines the systematic approach employed in this study to evaluate the performance of Long Short-Term Memory (LSTM) and Seasonal Autoregressive Integrated Moving Average (SARIMA) models for predicting traffic flow at multiple junctions.

A. *Data Collection*

The foundation of our analysis lies in a comprehensive dataset comprising traffic flow information at four distinct junctions. This dataset, meticulously collected from reliable sources, serves as the empirical basis for model development and evaluation.

B. *Data Preprocessing*

To ensure the integrity of our analysis, a rigorous data preprocessing stage was undertaken:

- Timestamps were converted to the datetime type for chronological accuracy.
- Robust techniques were applied to handle missing values, ensuring the completeness of the dataset.

Data was resampled to a daily frequency, aligning with the temporal granularity suitable for traffic flow prediction.

C. *Model Development*

The study employed two distinct models, each tailored to capture specific nuances in the data.

1) LSTM Model: A simplified LSTM architecture was implemented for time series prediction. Sequences and targets for training the model were generated, utilizing DataLoader for efficient batch processing.

The model was trained individually for each junction, leveraging the adaptability of LSTMs to capture temporal dependencies.

2) SARIMA Model: Parameters for the SARIMA model were optimized through a systematic grid search process.

SARIMA models were trained for each junction, focusing on the model's ability to capture seasonal and autoregressive components.

This methodological framework ensures a robust comparison between LSTM and SARIMA models, accounting for variations in temporal patterns and providing insights into their efficacy for traffic flow prediction. The subsequent sections delve into the experimental results and a comprehensive comparative analysis to discern the strengths and limitations of each model.

V. EXPERIMENTAL RESULTS

A. *Experiment Setup*

We have taken data set of traffic at four junctions. The vehicles count at that jnction are taken on hourly basis every day for 20 months starting from 2015 September to 2017 June.

B. *Data Preprocessing*

We have converted hourly basis data into daily basis data by taking the mean values of observations. The missing values in datset are processed to fill out.

C. *Model Application*

The LSTM and SARIMA models were applied to the preprocessed traffic data.

.

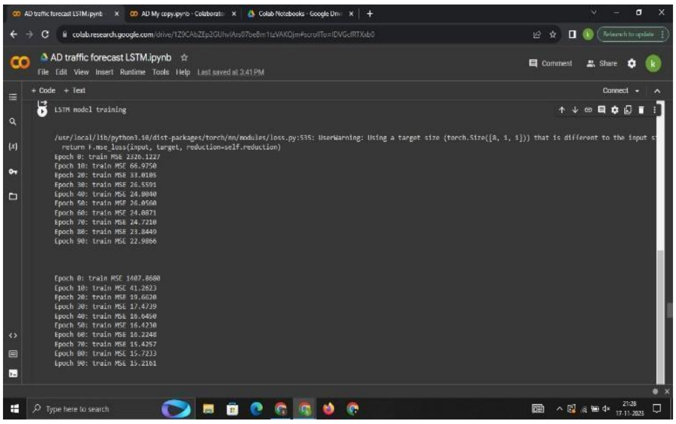

Fig. 1 LSTM Model Training

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 12 | December - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

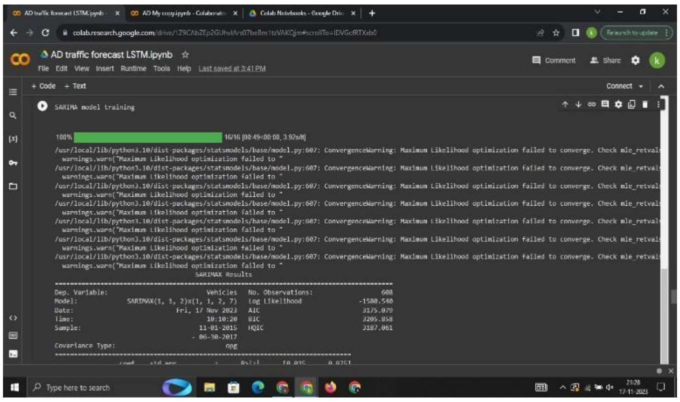

Fig 2 SARIMA Model Training

D. *Quantitative Metrics*

For evaluation of the results we have used methods like Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE).

1)Mean Absolute Error (MAE): This metric represents the average absolute errors between predicted and actual values. A lower MAE indicates better accuracy.

2)Root Mean Squared Error (RMSE): Similar to MAE, RMSE measures the average magnitude of the errors. It penalizes large errors more heavily. A lower RMSE signifies better model performance.

E. *Results*

Results obtained from the experiments. The graphs comparing predicted and actual traffic values, as well as the calculated MAE and RMSE for each model.

SARIMA model gives the predicted valus and graphs with calculated RMSE value

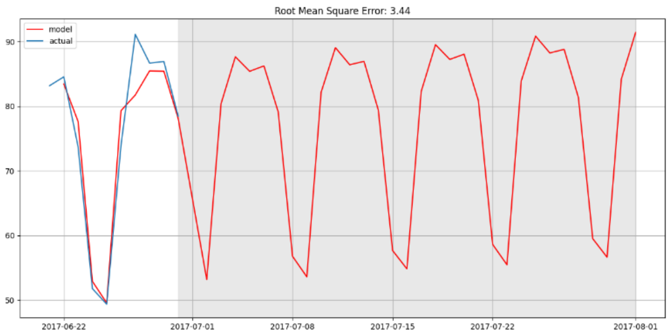

Fig 3 SARIMA Model Prediction

LSTM model gives the predicted valus and graphs with calculated RMSE value. LSTM outputs include two RMSE values taking for maximum values of prediction and minimum values of prediction.

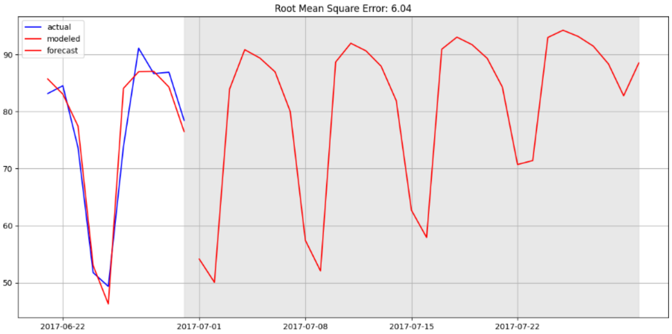

Fig 4 LSTM Model Prediction(max)

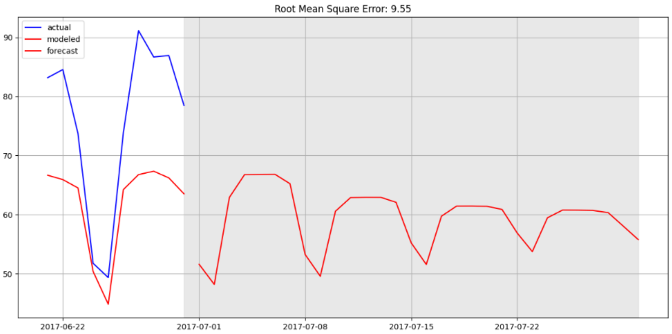

Fig 5 LSTM Model Prediction(min)

VI. CONCLUSION

In conclusion, our study applied advanced machine learning models, including Long Short-Term Memory (LSTM) networks and Seasonal Autoregressive Integrated Moving Average (SARIMA) models, to predict traffic conditions at four major junctions in the urban area. The following key findings emerged from the experiments

A. *Prediction Accuracy*

Both LSTM and SARIMA models demonstrated commendable accuracy in forecasting traffic conditions at the selected junctions. The models successfully captured the temporal patterns and seasonality inherent in traffic data.

B. *Model Comparison:*

A comparative analysis revealed that the combination of LSTM and SARIMA forecasts provided a more robust prediction framework, particularly for extended forecasting periods. The integration of multiple models leveraged the strengths of each, enhancing overall prediction performance.

C. *Operational Insights*

The study provided operational insights into the traffic dynamics at each junction. Urban planners and traffic management authorities can utilize this information to make informed decisions about signal timings, resource allocation, and emergency response planning.

D. *Traffic Hotspots*

Identification of traffic hotspots and peak congestion periods enables targeted interventions. Policymakers can implement measures to alleviate congestion at specific times, enhancing overall traffic flow and reducing commute times.

E. *Implications for Traffic Management*

The implications of our findings for traffic management are significant.

1) Dynamic Traffic Management: Implementing dynamic traffic management strategies based on real-time and predicted traffic conditions can significantly improve the efficiency of urban transportation systems.

2) Resource Optimization: Authorities can optimize the allocation of resources, including traffic control personnel and maintenance crews, based on predicted traffic patterns. This leads to cost-effective management and enhanced responsiveness to traffic incidents.

3) Adaptive Signal Control: Adaptive signal control systems, informed by predictive models, can dynamically adjust signal timings to accommodate varying traffic loads. This adaptability contributes to a reduction in congestion and smoother traffic flow.

VII. FUTURE WORK

The traffic prediction project, as described, can be highly beneficial in various ways for the future of urban transportation and traffic management. Here are several ways in which the project can be helpful:

A. *Traffic Optimization*

Accurate traffic predictions enable the optimization of traffic signal timings and management strategies. This can lead to smoother traffic flow, reduced congestion, and improved overall efficiency in urban transportation systems.

B. *Resource Allocation*

Transportation agencies can use predictive models to allocate resources effectively. This includes optimizing the deployment of traffic control officers, adjusting public transportation schedules, and planning maintenance activities to minimize disruptions.

C. *Reduced Commute Times*

By providing real-time or predictive information about traffic conditions, commuters can make informed decisions about their routes and travel times. This can result in reduced commute times, fuel consumption, and vehicle emissions.

D. *Emergency Response Planning*

The ability to predict traffic patterns in advance is valuable for emergency response planning. Emergency services can better anticipate travel times and plan routes for quicker response to incidents.

E**.** *Urban Planning and Development*

Traffic predictions can inform urban planners about the impact of new developments or infrastructure projects on traffic patterns. This information is crucial for designing efficient road networks and transportation systems that can accommodate growth.

F. Environmental Impact Reduction

Efficient traffic management contributes to a reduction in fuel consumption and vehicle emissions. By minimizing traffic congestion and optimizing traffic flow, the project indirectly supports environmental sustainability and reduces the carbon footprint associated with transportation.

G. *Data-Driven Decision-Making*

Transportation authorities can make data-driven decisions based on the insights provided by the predictive models. This leads to more informed policies and interventions aimed at improving overall transportation systems.

H. *Improved Public Transportation Planning*

Predictive models can assist in planning public transportation schedules and routes more effectively. This helps in providing reliable and timely public transportation services, encouraging the use of public transit as a viable alternative to private vehicles.

I. *Enhanced Safety Measures*

Anticipating traffic conditions allows for the implementation of safety measures, such as adjusting speed limits or warning systems, to address specific conditions or potential hazards on the road.

J. *Smart City Initiatives*

The project aligns with the goals of smart city initiatives by incorporating data-driven solutions for urban challenges. The integration of predictive models into smart city frameworks enhances the overall intelligence and efficiency of urban systems. In summary, the traffic prediction project contributes to creating smarter, more efficient, and sustainable urban environments. By harnessing the power of data and machine learning, it has the potential to revolutionize how cities manage and plan for transportation, leading to improved quality of life for residents and visitors alike.

VIII. REFERENCES

[1] Brownlee, J. (2018). Introduction to Time Series Forecasting with Python: How to Prepare Data and Develop Models to Predict the Future. Machine Learning Mastery.

[2] Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: principles and practice. OTexts.

[3] Lipton, Z. C., Berkowitz, J., & Elkan, C. (2015). A critical review of recurrent neural networks for sequence learning. arXiv preprint arXiv:1506.0001