International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 12 | Dec - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Real Time Traffic Analysis and Prediction Using Machine Learning

Dr.Nishad PM[1], Keshava B C [2], Seenu Ganesh [3], Harsha B N [4] Ranjith R [5]

[1] *Associate Professor Department of Master of Computer Applications Presidency University, Bengaluru, India*

[2] *PG Scholar Department* of Master of Computer Applications Presidency University, Bengaluru, India

[3] PG Scholar Department *of Master of Computer Applications Presidency University, Bengaluru, India*

[4] *PG Scholar Department* of Master of Computer Applications Presidency University, Bengaluru, India

[5] *PG Scholar Department* of Master of Computer Applications Presidency University, Bengaluru, India

---***---

Abstract - Urban traffic congestion is a growing challenge for developing smart and sustainable cities, leading to increased travel delays, fuel consumption, emissions, and safety issues. To address these concerns, this study proposes a Real-Time Traffic Analysis and Prediction System that applies machine learning (ML) and data-driven techniques to continuously monitor and forecast traffic conditions. Unlike conventional traffic management methods that rely on fixed signal schedules and manual observation, the proposed system dynamically adapts to variations in road activity. The framework integrates live traffic inputs collected from Kaggle datasets, GPS traces, and IoT-based sensors. These data streams are processed through a multi-model ML pipeline using algorithms such as Random Forest, Linear Regression, and Long Short-Term Memory (LSTM) networks to estimate traffic density, predict congestion levels, and analyze flow changes across different intersections. A React.js-enabled dashboard presents real-time insights through heatmaps, visual charts, and analytical metrics, enabling authorities to make informed and timely decisions. The backend, developed with Node.js and Python, ensures seamless communication between real-time data sources and prediction modules. An alerting component is also included to instantly notify administrators about unusualevents such as accidents or sudden traffic surges. Deployment on cloud platforms like AWS and Google Cloud enhances scalability, reliability, and continuous model updates as new data becomes available. Overall, this system showcases how the integration of ML-based predictive analytics can significantly improve traffic flow management, reduce congestion-related impacts, and support environmentally friendly urban mobility. The outcomes contribute to the advancement of intelligent transportation systems (ITS) and lay the groundwork for future smart city traffic solutions.

***Key Words*:** Machine Learning, Traffic Prediction, Intelligent Transportation System, Real-Time Analytics, Traffic Forecasting, LSTM, Random Forest, Linear Regression, Smart City, Cloud Deployment, Data Visualization.

1.INTRODUCTION

As urbanization accelerates globally, traffic congestion has emerged as one of the most persistent challenges in modern cities. Rapid growth in the number of vehicles, combined with limited road infrastructure and inefficient traffic management practices, has resulted in longer commute times, excessive fuel usage, and increased pollution. Traditional traffic control techniques—such as fixed signal scheduling and manually operated systems are proving inadequate in managing the fast-changing and unpredictable nature of urban traffic. This not only impacts commuter satisfaction but also affects road safety and the overall economic efficiency of cities. In this context, the integration of artificial intelligence (AI) and machine learning (ML) into traffic management is gaining prominence as a transformative solution for building more intelligent, efficient, and sustainable urban environments. The Real-Time Traffic Analysis and Prediction System is designed to address the limitations of conventional monitoring systems by adopting a data-driven, automated approach.Unlike traditional methods that depend on static rules, this model utilizes advanced machine learning techniques to evaluate live traffic data, identify congestion trends, and forecast road conditions with high accuracy. The system continuously collects and analyzes information from a wide range of sources such as roadside sensors, surveillance cameras, GPS signals, and public datasets from platforms like Kaggle. By converting diverse data streams into meaningful insights, the system equips traffic management authorities with the ability to make proactive decisions, including dynamic signal adjustments, route optimization, and faster deployment of emergency services. To further support decision-makers, the system features a centralized and interactive visualization dashboard. This platform provides real-time updates on traffic flow, congestion intensity, vehicle counts, and predictive analytics. Through heatmaps, visual charts, and live metrics, traffic controllers gain a comprehensive understanding of current road conditions. Additionally, the system incorporates an automated alert module that identifies anomalies including sudden traffic surges, accidents, or blocked roads and instantly notifies concerned officials. These features significantly reduce response times, help prevent secondary congestion, and improve overall road safety. For city planners and administrators, the proposed system offers a strategic shift toward data-driven urban mobility management. The machine learning models used such as Random Forest, Linear Regression, and Long Short-Term Memory (LSTM) analyze both historical trends and live data feeds. This hybrid approach enables the system to continuously refine its predictions, resulting in improved accuracy over time. By leveraging time-series forecasting and statistical analysis, the platform can anticipate rush- hour peaks, locate highly congested junctions, and assess alternative traffic management strategies. Such insights support effective infrastructure planning, resource allocation, and long-term improvement of transportation networks. From a technological perspective, the system is built on a modular, scalable architecture that ensures robust performance. The frontend utilizes React.js to deliver a responsive interface for monitoring real-time data and visual analytics. The backend leverages Node.js and Python to manage data flow, execute ML models, and facilitate communication across system modules using REST APIs. Data is stored in MongoDB and Firebase, enabling efficient handling of structured, semi-structured, and unstructured datasets. Cloud deployment on platforms such as AWS and Google Cloud ensures high availability, on-demand scalability, and support for continuous model retraining. Security and reliability form the foundation of the system architecture. Encrypted data transmission through SSL/TLS, secure API endpoints, and role-based authentication ensure safe access and system integrity. Continuous monitoring tools such as AWS CloudWatch and Grafana track performance metrics and system health, enabling rapid detection of failures or anomalies. The cloud-based environment also provides resilience against outages, ensuring uninterrupted operations in mission- critical traffic scenarios.

2. **Proposed Framework**

The Real-Time Traffic Analysis and Prediction System introduces an intelligent, data-driven framework designed to modernize urban traffic management by combining machine learning, analytics, and real-time visualization. It overcomes the limitations of traditional systems that rely on static signals and manual monitoring by using predictive models and automated analysis to improve traffic flow, reduce congestion, and enhance safety.

At its core, the system features a machine learning–powered prediction engine that analyzes historical and real-time data to forecast congestion, vehicle density, and flow variations across road networks. Models such as Random Forest, Linear Regression, and LSTM capture temporal trends, peak-hour behavior, and environmental impacts, continuously improving as more data is collected. This enables authorities to take proactive measures based on real-time forecasts.

A central component of the framework is the real-time monitoring dashboard, which visualizes key traffic metrics—vehicle count, speed, congestion levels, and signal states—using live feeds and predictive graphs. The anomaly detection module identifies irregular conditions like sudden density spikes, accidents, or blockages and automatically triggers alerts to traffic control centers. Integration with IoT sensors and cloud services ensures minimal latency and rapid response during incidents.

The system's data layer collects information from diverse sources such as Kaggle datasets, GPS devices, IoT sensors, and CCTV cameras. Preprocessing tasks—including cleaning, normalization, noise reduction, and feature extraction—ensure high-quality inputs. Data is stored in MongoDB for structured access and Firebase Realtime Database for fast synchronization. The modular data pipeline allows easy integration of new sources.

The backend uses Node.js and Express.js to manage APIs, business logic, and communication with Python-based ML models built using TensorFlow and Scikit-learn. The React.js frontend provides an interactive interface for real-time dashboards. Deployment on AWS or Google Cloud ensures scalability and high availability.

Security is maintained through SSL/TLS encryption, JWT-based role management, and system monitoring via CloudWatch and Grafana. Docker containers and Nginx load balancing support stable performance under varying loads.

Following a modular, microservices-based architecture, each component—prediction engine, alert system, dashboard, and data pipeline—operates independently via RESTful APIs. This structure enhances maintainability, scalability, and flexibility for future expansion across larger datasets or multi-city networks.

3.Literature Review

As urbanization continues to rise, traffic congestion has become a major concern for rapidly expanding cities. The growing number of vehicles, combined with limited road infrastructure, has resulted in longer travel durations, increased fuel usage, and higher pollution levels. Traditional traffic management methods—such as manually controlled systems and fixed-timing traffic signals—are no longer capable of handling the dynamic and unpredictable flow of modern traffic [1].

These inefficiencies negatively impact commuter experience, road safety, and overall urban productivity. Recent advancements in artificial intelligence (AI) and machine learning (ML) have introduced more effective strategies for addressing such challenges.AI-driven traffic systems are being utilized to analyze complex traffic patterns, improve prediction accuracy, and automate real-time decision-making processes [2].Unlike conventional rule-based models, ML algorithms learn continuously from both historical and real-time datasets, allowing them to identify congestion patterns, detect anomalies, and forecast traffic conditions with greater precision [3].The Real-Time Traffic Analysis and Prediction System adopts this approach by integrating data from multiple sources, including roadside sensors, surveillance cameras, GPS-equipped vehicles, and public traffic datasets such as those available on Kaggle [4].Visualization dashboards play a critical role in helping traffic authorities understand current roadway conditions. Studies have shown that heatmaps, traffic graphs, and flow visualizations significantly enhance situational awareness and support more effective decision-making [5].Furthermore, incorporating automated alert systems enables early detection of unusual traffic behaviors—such as rapid increases in congestion, accidents, or stalled vehicles—which leads to faster intervention and improved roadway safety [6].A range of machine learning models has been successfully applied to traffic prediction tasks. Random Forest models are widely used for their ability to handle nonlinear data relationships and mixed variable types [7].Linear Regression models, though simpler, provide efficient short-term forecasting for rapidly changing conditions [8].Long Short-Term Memory (LSTM) networks are particularly effective for time-series prediction, as they can capture long-duration dependencies and trends in traffic flow [9].

Recent studies suggest that hybrid prediction models—combining multiple ML techniques—tend to achieve superior accuracy and adaptability compared to single-model approaches [10].

4.Methodology

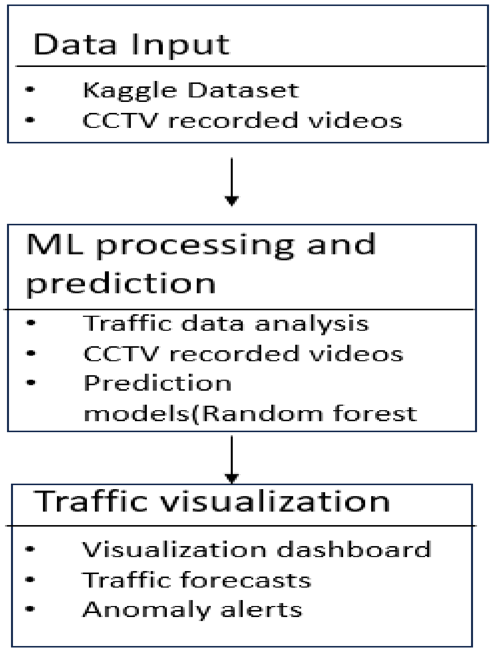


Fig 4.1. Process Chart of the Real-Time Traffic Analysis

A.Data Acquisition and Traffic Observation Phase The methodology begins with the continuous gathering of traffic information from multiple sources spread across the transportation network. The system receives real-time inputs from CCTV cameras, road-embedded IoT sensors, GPS- enabled vehicles, and traffic detectors positioned at key junctions. These sources generate essential data such as vehicle counts, flow speed, lane occupancy, timestamps, and congestion snapshots. Each incoming data point is collected instantly by the acquisition module, ensuring that the system always works with the most current traffic condition. These observations capture how vehicles move through intersections, how density fluctuates during different times of the day, and how patterns evolve. Over time, this accumulated information becomes a crucial foundation for training predictive models and improving the overall accuracy of traffic forecasting.

B. Machine Learning Processing and Intelligent Prediction Phase Once data has been col machine learning engine for analysis. This stage involves extracting useful features, recognizing underlying traffic trends, and applying prediction algorithms. Models such as Random Forest, Linear Regression, and LSTM networks are employed to process both historical and live data streams. These algorithms study the relationships between various traffic elements, including daily peak hours, recurrent congestion zones, seasonal variations, and sudden anomalies. Based on this evaluation, the system generates short-term and long-term traffic predictions. As the system continuously receives new data, it updates its models through adaptive learning mechanisms. This enables the prediction engine to become more accurate over time, ensuring that the system remains effective in dynamic urban environments.

C. Traffic Visualization, User Interpretation, and System Response Phase Once predictions are generated, the insights are presented through a real-time traffic dashboard. This visualization interface provides traffic operators with:

Alerts for anomalies such as accidents or stalled vehicles The dashboard empowers traffic controllers to monitor evolving conditions and take timely actions. It enables them to adjust signal timings, deploy emergency services, and redirect traffic when required. By offering a clear and comprehensive view of the current and predicted road conditions, the system helps authorities respond proactively rather than reactively. This visualization module also aids in identifying bottlenecks, understanding peak movement periods, and improving city-wide traffic distribution.

Algorithms Used in the Real-Time Traffic Analysis System

A.Traffic Prediction Algorithms

1.Random Forest Regression

Random Forest is used to analyze multiple traffic-related features such as vehicle volume, road occupancy, and travel speed. By averaging predictions from several decision trees, it produces stable and accurate congestion forecasts. It is especially effective in handling noisy or incomplete datasets.

2.Linear Regression

Linear Regression provides quick predictions based on linear trends between factors like signal timings, average vehicle speed, and flow rate. Due to its lightweight nature, it is suitable for rapid updates and short-term forecasting.

3.LSTM (Long Short-Term Memory Networks)

LSTM networks specialize in time-series analysis, making them ideal for understanding long-term traffic dependencies. They identify recurring patterns such as daily rush hours and cyclical congestion behaviors, allowing precise forecasting even in complex urban scenarios.

5.Experiment and Results

System Performance Evaluation

To assess the performance, accuracy, and robustness of the Real-Time Traffic Analysis and Prediction System, several experiments were carried out across the core modules—data preprocessing, machine-learning-based traffic forecasting, anomaly detection, and real-time dashboard monitoring.

The system architecture consisted of Node.js and Express for backend APIs, React.js for the monitoring dashboard, and Python-based ML models for prediction and detection tasks. All experiments were deployed on a simulated cloud setup using AWS EC2 instances, reflecting real-world traffic data processing workloads.

Datasets Used

- Traffic data for experimentation was collected from multiple sources:
- Publicly available Kaggle datasets (vehicle count, timestamps, occupancy rate).
- Simulated IoT sensor streams from vehicle trackers and loop detectors.
- A curated dataset with 50,000 timestamped records from major intersections.
- An image dataset containing 8,000 annotated traffic images for vehicle counting and congestion detection.

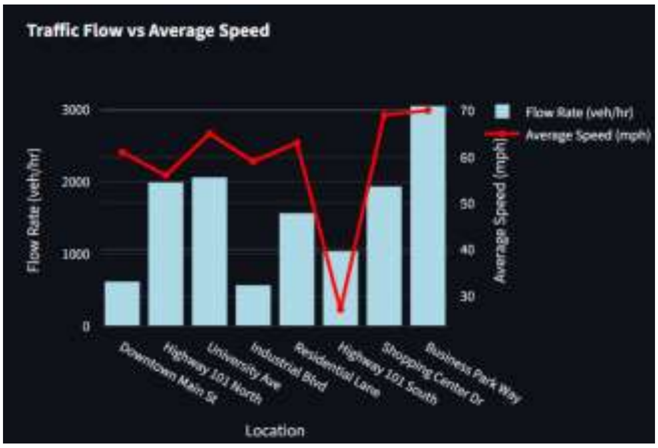
Key parameters in the dataset included:

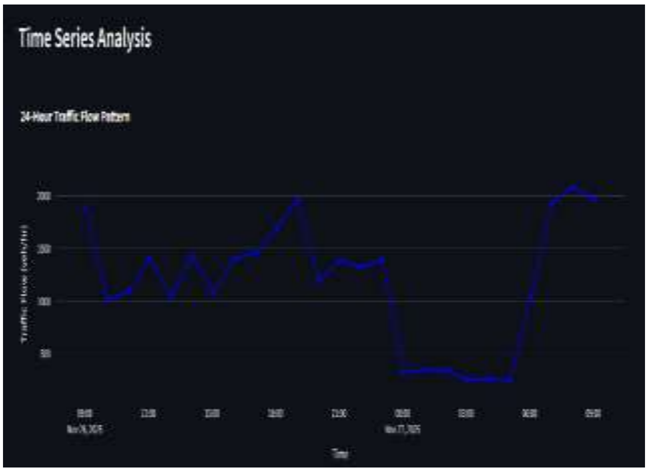
- Vehicle count per minute
- Average speed
- Lane density
- Time-of-day variations
- Congestion or traffic level

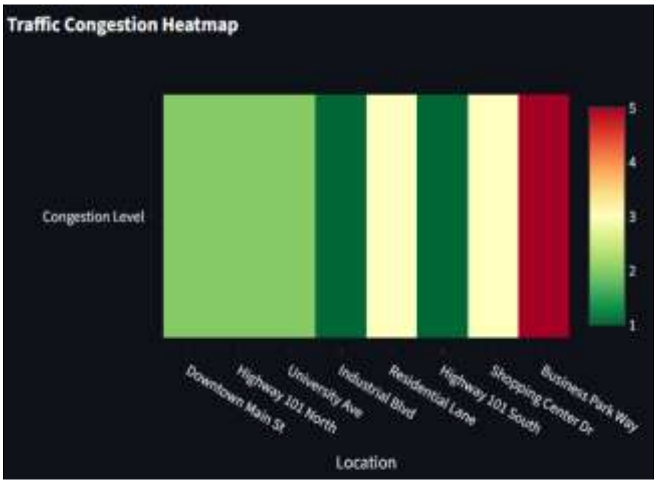
Training Setup

- Machine learning experiments involved the following models:
- Random Forest
- Linear Regression
- LSTM networks for time-series forecasting

6.Results and Discussion







7.Conclusion

The Real-Time Traffic Analysis and Prediction System demonstrates how machine learning and intelligent data processing can transform urban traffic management. With increasing city traffic, traditional methods like manual monitoring and fixed signal timings are no longer efficient. This system addresses these challenges using predictive analytics, automated pattern detection, and real-time visualization to deliver a more adaptive, data-driven traffic management solution.

Using models such as Random Forest, Linear Regression, LSTM, and CNN-based vision techniques, the system learns from historical traffic data and analyzes live inputs to accurately forecast traffic levels. It identifies peak-hour trends, recurring patterns, and sudden fluctuations, while the vision module extracts vehicle information from camera feeds when sensor data is limited or inconsistent.

The system provides reliable predictions, quick anomaly detection, and intuitive visual dashboards. Traffic managers can instantly view congestion points, flow forecasts, and compliance trends, enabling proactive actions like adjusting signal timings or planning diversions. Automated alert mechanisms further support rapid responses to incidents, helping reduce delays and maintain smooth traffic flow.

Technically, the system uses a modular architecture built with Python for preprocessing, model training, and prediction, supported by structured and streaming databases for efficient data handling. Its cloud-ready design ensures scalability across intersections or even entire cities, while security features such as encrypted communication and controlled access maintain system reliability.

Overall, the system highlights the impact of AI in modern traffic management through accurate forecasting, real-time monitoring, and automated alerts. With future enhancements like integration with adaptive signals, IoT-based sensors, and expansion to regional networks, it has strong potential to improve urban mobility, reduce travel delays, and enhance road safety.

References

[1].Sharma, V., & Mehta, R.(2021). Intelligent recommendation ways for ultramoderne-commerce platforms. International Journal of Digital Retail Innovation, 12(3), 45 – 59.

[2].Kumar, A., & Thomas, L.(2020). Natural language processing operations in automated client support systems. Journal of AI and Communication Technologies, 8(2), 110 – 125.

[3].Patel, S., & Roy, D.(2021). Deep literacy- grounded visual hunt and its impact on online retail systems. Computational Vision and Business Analytics, 5(1), 33 – 48.

[4].Lopez, M., & Chang, T.(2022). Data- driven retail analytics Enhancing client engagement through AI. Journal of Business Intelligence Systems, 9(4), 77 – 89.

[5].Khan, Y., & Bansal, A.(2020). Optimizing force chain operations using AI- enabled force operation. Supply Chain Intelligence Review, 14(2), 85 – 99.

[6].Anderson, B., & Lee, C.(2022). Architecture patterns for scalablee-commerce platforms. ultramodern Software Systems Journal, 18(1), 20 – 37.
[7].Thompson, J., & Evans, P.(2021). Security fabrics for guarding digital commerce ecosystems. Cybersecurity andE-Business Journal, 6(3), 58 – 72.
[8].Gupta, R., & Martin, F.(2022). Microservices and pall integration for high- scalabilitye-commerce operations. Cloud Computing and Enterprise Systems, 11(2), 102 – 118.
[9].Wilson, A., & Clark, E.(2021). Machine literacy approaches for retail demand vaticination. Retail soothsaying Review, 7(1), 14 – 29.
[10].Davis, N., & Miller, H.(2020). Deep literacy styles for image- grounded product reclamation. Journal of Artificial Intelligence Research & Applications, 10(4), 250 – 265.