International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 10 | Oct - 2026 SJIF Rating: 8.659 ISSN: 2582-3930

Big Data Analytics for Smart Cities

Mrs. Gadekar Kiran Satyawan,

Mrs. Asmita Mohan Phatangare

Department of Master of Computer Application,
JCEI's Jaihind Institute of Management and Research, Vadgaon Sahani

Email: kirangadekar241@gemail.com,

asmitafatangare41@gmail.com

Abstract - Urban growth and the increasing number of vehicles have led to various transportation problems. These include road congestion, travel delays, fuel wastage and air pollution. Conventional traffic control approaches often struggle to manage continuously changing road conditions. This project explores the use of large-scale data analysis and machine learning to improve traffic monitoring and transportation planning. Information gathered from GPS-enabled devices, roadside sensors and monitoring cameras can be processed using distributed computing technologies such as Apache Hadoop and Apache Spark. Artificial intelligence techniques can help identify traffic patterns estimate congestion levels. Support dynamic traffic signal control. In addition, visual data representations and route analysis can assist in locating roads and identifying alternative travel paths. The suggested system focuses on improving traffic flow, cutting travel delays and using existing road infrastructure effectively. Overall combining data-processing methods with intelligent traffic analysis can support the development of reliable and sustainable urban transportation systems.

***Key Words*:** Intelligent Traffic Control, Large-Scale Data Analysis, Artificial Intelligence, Live Traffic Monitoring, Distributed Computing, Traffic Flow Forecasting, Road Congestion Analysis, Visual Data Representation, Smart Transportation Networks, Dynamic Signal Management.

1. INTRODUCTION

The development of cities needs an effective transportation system that can handle the growing flow of vehicles and people. I notice that the quick rise of cities is making more people decide to live inside cities. More people mean private cars on the roads. Traffic congestion causes longer travel times, fuel use, more air pollution and hard times for everyday commuters. These problems also hurt activity and the overall quality of city life. Therefore, new technologies must improve traffic monitoring, cut road congestion and make transportation systems more reliable and efficient. Big data analytics and real- time data processing give ways to handle these transportation challenges. Traffic information can be gathered from road sensors, surveillance cameras, GPS devices and other monitoring systems. By studying this information traffic authorities can spot congestion- prone areas, track vehicle movement check road conditions and learn traffic patterns at times of the day. Machine learning techniques can also help predict traffic conditions from past and present data. In addition, analytical results can aid traffic signal timing, route planning and quicker responses to road accidents or unexpected traffic disruptions. The integration of data analytics, real- time processing and smart traffic control can lead to better urban transportation management. These methods can shorten delays at traffic signals, distribute vehicles more efficiently across roads, and

improve traffic flow. Proper traffic management can also reduce fuel consumption and minimize harmful emissions from vehicles. The success of these solutions depends on the quality of the collected data, reliable communication systems and precise analysis. This project focuses on applying data technologies to traffic monitoring and congestion analysis with the goal of supporting smoother traffic flow, safer roads and more sustainable transportation, in smart cities.

1.1 Problem Statement

I see that rapid urbanization and more cars on the road cause traffic jams to make trips longer and worsen transportation and the traffic we have the slower we all move. old ways of controlling traffic have a time dealing with changing road situations. To fix these problems big data analytics, real-time data processing and machine learning can be used to watch traffic patterns guess where traffic will get bad make traffic lights work better and find ways to go. The main goal is to create a traffic management system that makes traffic move better cuts down on delays and makes transportation work better.

1.2 Objectives

• To study urban traffic data and its features with the help of big data concepts.

• To optimize the traffic signals and routing to reduce the delays.

• To improve the traffic flow and decrease the time taken.

• To develop an efficient and effective system for intelligent traffic management.

1.3 Hypothesis

• H_0: (Null Hypothesis): Big data analytics and real-time processing do not have a significant improvement effect on urban traffic management.

• H_1: (Alternative Hypothesis): Big data analysis and real-time processing can improve traffic flow by reducing congestion and travel delays.

2. LITERATURE REVIEW

The rise of cities has made traffic monitoring and transportation management more important. Researchers are looking at technologies to make road traffic smoother and cut delays. Data from road sensors GPS devices, cameras and other digital sources helps us study how vehicles move and spot traffic congestion. Big data processing frameworks like Apache Hadoop and Apache Spark are capable of managing volumes of transportation information. These tools are designed to process datasets efficiently. They can handle data from sources related to transportation. This includes traffic data, GPS signals and vehicle tracking information. These tools are designed to process datasets efficiently. They can handle data from sources related to transportation. This includes traffic data, GPS signals and vehicle tracking information. Processing transportation data, at scale becomes possible with these technologies. They help organizations make decisions based on accurate and timely information. These technologies help analyze traffic patterns and spot changes in road conditions. Machine learning techniques can also take in traffic data to forecast how busy the roads will be and help with planning, for traffic. Smart traffic lights are another way to make traveling on the roads better.

Intelligent traffic signals provide another way to improve road transportation. These systems can change signal timings as traffic conditions change. Also, traffic analysis tools can point out congested routes improve travel planning and cut down on unnecessary waiting times. Putting intelligent traffic management systems into practice has many challenges. Challenges include keeping data quality high processing information handling implementation costs and merging data from different sources. Systems must also react quickly to

changes in traffic conditions. Given these points combining data technologies with machine learning and real- time traffic monitoring can offer an effective way to manage modern transportation. This blend can improve traffic flow, cut travel delays and help build sustainable cities.



3. METHODOLOGY

This paper wants to look into the role of data analytics and real-time processing in managing traffic in smart cities by using sensors, GPS and machine learning. Case study methodology is used in this study to check traffic management and control by looking at traffic data and GPS information. Data analytics, including machine learning and real-time processing are used to study traffic management.

3.1 Data Collection

Traffic information is gathered from sources such as

1. Traffic Sensors: To record the number of vehicles, speed and traffic congestion.
2. GPS Data: Used to determine vehicle locations and monitor their direction of movement.
3. Surveillance Cameras: To monitor the flow of traffic and detect congestion or accidents.
4. Mobile Applications: To provide traffic information and collect feedback from drivers.

3.2 Big Data Analytics Process

• Apache Hadoop: To store and manage amounts of traffic data.

• Apache Spark: To perform fast data processing, analysis and machine learning on traffic data.

• Machine Learning: To analyze traffic patterns predict congestion and study traffic behaviour.

3.3 Tools and Technologies

• Apache Hadoop: Used to store and process amounts of traffic data.

• Apache Spark: Helps process volumes of data quickly and runs machine learning tasks efficiently.

• Apache Kafka: Used to collect and stream real-time traffic data without delays.

• Apache Flink: Processes data continuously. Handles real-time data flow effectively.

• Python: Used to analyze data run machine learning models and create visualizations for understanding. • Machine Learning Algorithms: Applied to forecast traffic congestion and detect patterns in how vehicles move.

• GPS and Traffic Sensors: Gather data on vehicle locations, speeds and traffic density.

• Surveillance Cameras: Help monitor traffic flow and spot congestion or accidents as they happen.

• Database: Stores traffic records and manages up-, to-date traffic data efficiently.

4. Implementation

1. gather traffic data from sources such as sensors, cameras and GPS.
2. Save the collected information in a database for future use.
3. Clean the data by eliminating information.

4. I will analyze the data to understand the traffic conditions and look for patterns.

5. Highlight high-traffic areas.

6. Predict the traffic conditions.

7. I will adjust the traffic signals to control the flow of traffic and make sure commuters have a ride.

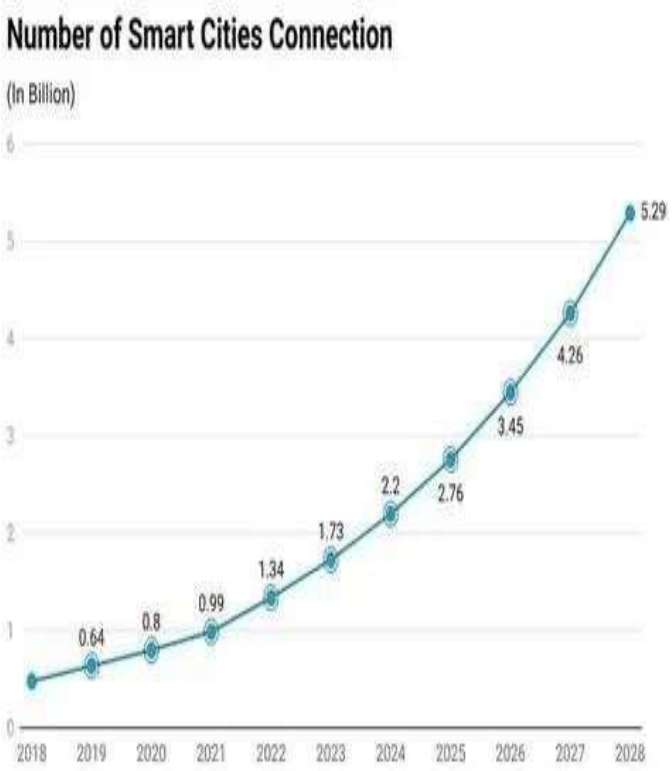
8. Recommend routes to avoid traffic.

9. Send alerts, about accidents and emergency services.

10. Display traffic information to the users and authorities.

11. Update the data continuously.

12. Monitor the system to improve the traffic conditions.



5. CONCLUSION

I believe that the traffic management system uses sensors, CCTV, GPS and other data sources to monitor traffic in areas. The traffic project can detect traffic jams, accidents and traffic congestion. The traffic system can control traffic by managing traffic lights and giving drivers routes. I believe that this study will promote traffic safety, cut traffic waiting time and optimize traffic flow. By doing it will reduce fuel waste and air pollution and it .

REFERENCES

[1] Soomro, K., Bhutta, M. N. M., Khan, Z., and Tahir, M. A. (2019). *Smart City Big Data Analytics: An Advanced Review.* Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery.
Link: https://doi.org/10.1002/widm.1319

[2] Karimi, Y., Haghi Kashani, M., and Akbari, M. (2021). *Leveraging Big Data in Smart Cities: A Systematic Review.* Concurrency and Computation: Practice and Experience.
Link: https://doi.org/10.1002/cpe.6379

[3] Cesario, E. (2023). *Big Data Analytics and Smart Cities: Applications, Challenges, and Opportunities.* Frontiers in Big Data.
Link: https://doi.org/10.3389/fdata.2023.1149402

[4] Abdelrahman, Y., Hajek, P., and Hikkerova, L. (2023). *Research Trends in the Application of Big Data in Smart Cities—A Literature Review.* Canadian Journal of Administrative Sciences.
Link: https://doi.org/10.1002/cjas.1734

[5] Atitallah, S. B., Driss, M., Boulila, W., and Ben Ghézala, H. (2020). *Leveraging Deep Learning and IoT Big Data Analytics to Support Smart Cities Development: Review and Future Directions.* Computer Science Review.
Link: https://doi.org/10.1016/j.cosrev.2020.100303

[6] Bibri, S. E. (2019). *On the Sustainability of Smart and Smarter Cities in the Era of Big Data: An Interdisciplinary and Transdisciplinary Literature Review.* Journal of Big Data.
Link: https://doi.org/10.1186/s40537-019-0182-7

[7] Olaniyi, O. O., Okunleye, O. J., and Olabanji, S. O. (2023). *Advancing Data-Driven Decision-Making in Smart Cities through Big Data Analytics: A Comprehensive Review of Existing Literature.* Current Journal of Applied Science and Technology.
Link: https://doi.org/10.9734/cjast/2023/v42i254181

[8] Gubareva, R., and Lopes, R. P. (2024). *Literature Review on the Smart City Resources Analysis with Big Data Methodologies.* SN Computer Science.
Link: https://doi.org/10.1007/s42979-023-02457-x

[9] Althabahi, M. H., Ahmad Jan, M., Brik, B., and Foufou, S. (2024). *Machine Learning-Based Big Data Analytics in Smart Cities: A Survey of Current Trends and Future Research Directions.* IEEE/ACM International Conference on Big Data Computing, Applications and Technologies.
Link: https://doi.org/10.1109/BDCAT63179.2024.00040

[10] *An Improved Large Scale Data Analytics for Smart Cities with Multimodal Data Fusion.* IEEE Conference Publication. Link: https://ieeexplore.ieee.org/abstract/document/10250451

[11] *Big Data Analytics for Smart Cities.* IEEE 2024 1st International Conference on Advances in Computing, Communication and Networking. Link: https://doi.org/10.1109/ICAC2N63387.2024.10895781

[12] Theodorakopoulos, L., and Theodoropoulou, A. (2026). *Big Data Analytics for Geospatial Decision-Making in Smart Cities: A Review of Spatial Data, GeoAI and Urban Digital Twins.* ISPRS International Journal of Geo-Information. Link: https://doi.org/10.3390/ijgi15070278

[13] Batty, M. (2013). *Big Data, Smart Cities and City Planning.* Dialogues in Human Geography. Link: https://doi.org/10.1177/2043820613513390

[14] Kitchin, R. (2014). *The Real-Time City? Big Data and Smart Urbanism.* GeoJournal. Link: https://doi.org/10.1007/s10708-013-9516-8

[15] Hashem, I. A. T., et al. (2016). *The Role of Big Data in Smart City.* International Journal of Information Management. Link: https://doi.org/10.1016/j.ijinfomgt.2016.05.002

[16] Zanella, A., Bui, N., Castellani, A., Vangelista, L., and Zorzi, M. (2014). *Internet of Things for Smart Cities.* IEEE Internet of Things Journal. Link: https://doi.org/10.1109/JIOT.2014.2306328

[17] Batty, M., et al. (2012). *Smart Cities of the Future.* The European Physical Journal Special Topics. Link: https://doi.org/10.1140/epjst/e2012-01703-3

[18] Neirotti, P., De Marco, A., Cagliano, A. C., Mangano, G., and Scorrano, F. (2014). *Current Trends in Smart City Initiatives: Some Stylised Facts.* Cities. Link: https://doi.org/10.1016/j.cities.2013.12.010

[19] Hashem, I. A. T., et al. (2015). *The Rise of Big Data on Cloud Computing: Review and Open Research Issues.* Information Systems. Link: https://doi.org/10.1016/j.is.2014.07.006

[20] Cesario, E., et al. (2016). *A New Approach to the Design of Smart City Applications.* Research on data-driven urban applications and intelligent city services. Link: https://scholar.google.com/scholar?q=Eugenio+Cesario+smart+city+applications+big+data