Traffic Congestion Prediction and Optimization

Using Python and Machine Learning

Developed By: Neha Palanati, Srinaini Garimella, Kommathoti Yanni Rohan

Smart Bridge Project Report - Applied data science

1. INTRODUCTION

1.1 OVERVIEW:

We are undertaking a project to leverage Machine Learning techniques in order to predict traffic congestion. We possess a dataset that captures the count of vehicles at various time intervals, enabling us to assess the level of traffic congestion. By analysing this data, we can identify the specific periods when traffic volume is at its peak. Armed with this knowledge, we can adapt our plans and make informed decisions that align with the prevailing traffic conditions, ensuring smoother travel experiences and efficient use of resources.

The primary objective of traffic prediction is to offer precise and timely information to support informed decision-making for drivers, transportation authorities, and stakeholders. By anticipating traffic patterns and congestion, it becomes possible to optimize routes, allocate resources efficiently, and enhance overall traffic management.

1.2 PURPOSE:

Considering recent developments regarding fuel scarcity and the increased awareness of environmental concerns, researchers have turned their attention towards exploring alternative fuels, hybrid vehicles, and other sustainable fuel options. This shift in focus stems from the historical abundance of petrol, which led to the construction of cars optimized for high-speed operation, comfort, and safety. As the automotive industry evolves, there is a growing need to address the challenges posed by fuel scarcity and reduce the excessive reliance on gasoline. One area of exploration involves redesigning the car's body to minimize aerodynamic losses. Spoilers are employed to mitigate these losses, necessitating the optimization of their shape.

This study aims to identify the most efficient spoiler design for automobiles, considering factors such as strength, aerodynamics, mass reduction, and fuel consumption. To evaluate changes in the aerodynamic properties of the spoiler's cross-section, a Computational Fluid Dynamics (CFD) analysis is conducted on a two-dimensional spoiler model. The findings of this analysis serve as a validation mechanism for understanding the impact of various design modifications on the spoiler's aerodynamic performance.

By leveraging advanced computational techniques and conducting rigorous analyses, this research endeavours to contribute to the development of optimal spoiler designs that enhance both the vehicle's aerodynamic efficiency and fuel consumption. These efforts align with the broader objective of fostering sustainable practices within the automotive industry and reducing the ecological footprint of automobiles.

2. LITERATURE SURVEY

2.1 EXISTING PROBLEM:

[1] The current challenge we face is the lack of knowledge about peak and low traffic times, leading to inconvenience during travel.

Traffic congestion is a prevalent issue characterized by the overabundance of vehicles on roads, leading to slower travel speeds, heightened fuel consumption, and increased pollution levels. This situation brings about inconvenience, delays, and reduced effectiveness for both commuters and transportation systems.

2.2 PROPOSED SOLUTION:

[2] The optimisation of traffic prediction involves leveraging machine learning models to assess traffic conditions at different times. By doing so, we can identify periods of low traffic and take appropriate measures, such as opting for alternate routes or implementing other strategies, to ensure smoother and more efficient travel experiences.

Data Collection: To collect relevant data, a comprehensive approach involves gathering information from diverse sources, including traffic cameras, sensors, GPS devices, and social media platforms. This data acquisition process aims to obtain real-time insights on traffic patterns, road conditions, and any incidents that may occur.

Traffic Prediction: To predict traffic patterns, machine learning techniques such as regression, time series analysis, or neural networks can be applied. By analysing both historical and real-time data, these ML models can make predictions about traffic congestion. These predictions incorporate variables such as the time of day, date, month, year, and special events to provide accurate forecasts.

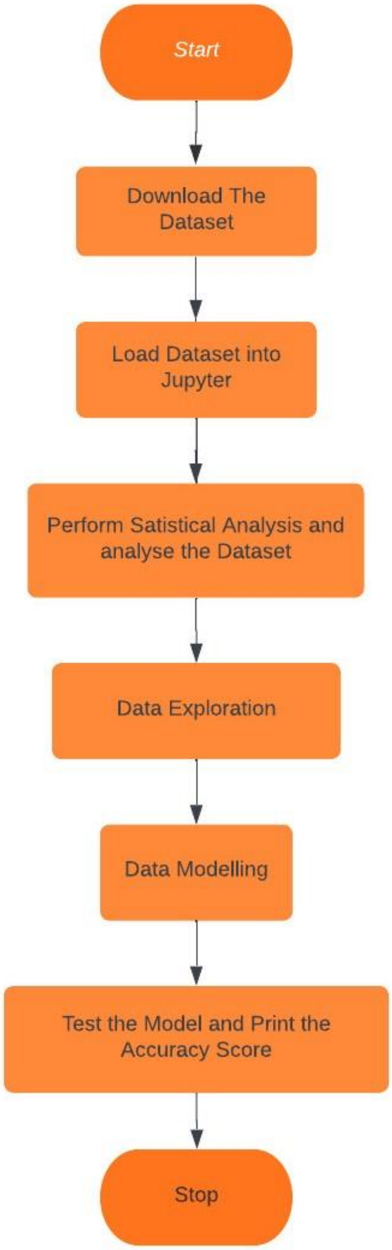
Optimal routing and traffic control: Based on the predicted traffic patterns, ML algorithms can suggest optimal routes for drivers to avoid congested areas. Additionally, ML can assist in traffic signal optimization and adaptive traffic control to dynamically adjust signal timings based on real-time traffic conditions.

Intelligent transportation systems: ML can be leveraged to develop intelligent transportation systems that can monitor traffic in real-time, identify traffic incidents, and automatically notify relevant authorities or emergency services.

Feedback loop: Continuously collect and analyse data on the effectiveness of the ML-based traffic management system. Use this feedback to improve the accuracy of traffic predictions and optimize traffic control strategies.

3. THEORITICAL ANALYSIS

3.1 BLOCK DIAGRAM:



3.2 HARDWARE/SOFTWARE DESIGNING:

1. The hardware requirements for traffic congestion prediction, some general recommendations for running the code and models include:
2. A computer with a modern processor (e.g., Intel Core i5 or higher) and at least 8GB of RAM.
3. A Python environment with the necessary packages installed, including NumPy, Pandas, Matplotlib, Scikit-learn, and Statsmodels.
4. A Jupyter Notebook environment for running the code and visualizing the results.
5. Sufficient storage space for storing the traffic congestion data and the trained models. It is worth noting that the hardware requirements may vary depending

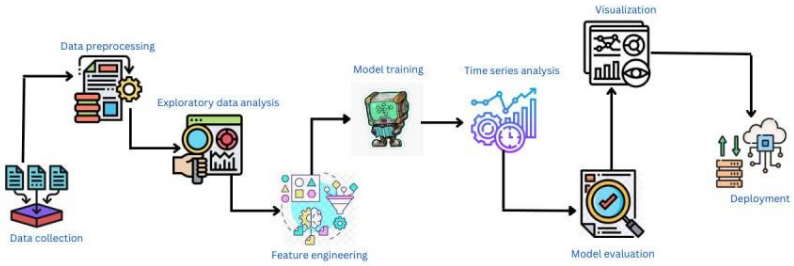
on the size of the traffic congestion dataset and the complexity of the models used. In general, it is recommended to use a computer with sufficient processing power and memory to ensure smooth and efficient execution of the code.

6. **HTML and CSS Tools:** Spyder itself provides an integrated web browser that can be used for this purpose. Alternatively, you can use other web browsers like Google Chrome, Mozilla Firefox, or Microsoft Edge to open and view your HTML and CSS files during the application development phase.

4. EXPERIMENTAL INVESTIGATIONS

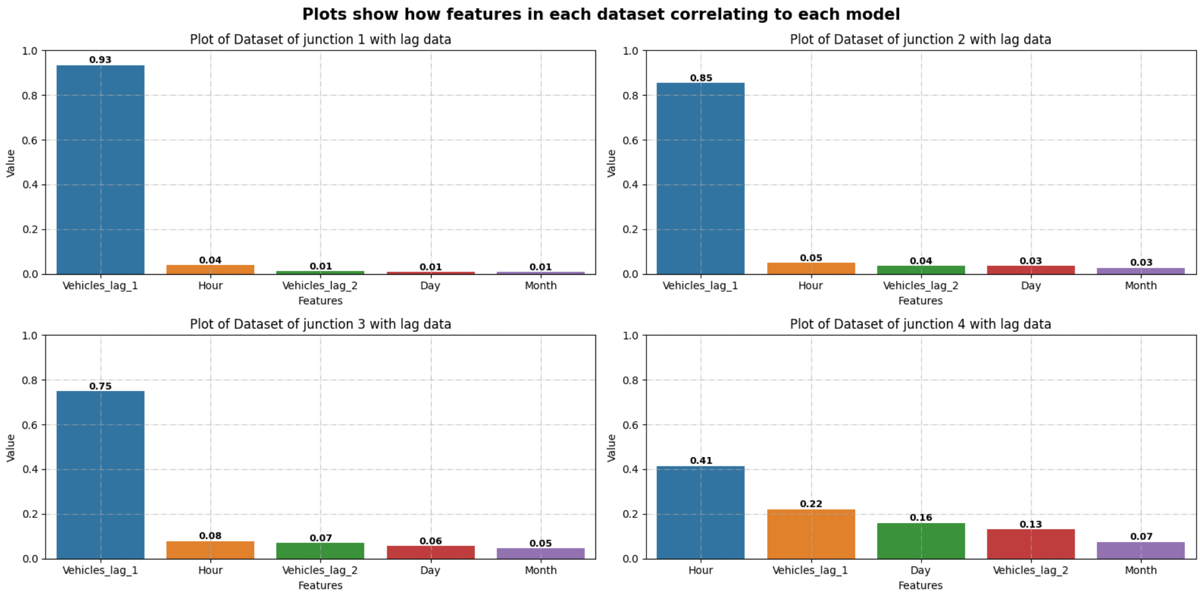
- The experimental investigations for this traffic congestion prediction involve applying various machine learning and time series analysis techniques to a dataset of traffic congestion data from Bangalore, India. The main steps involved in the analysis include:
- Data preprocessing and exploratory data analysis to understand the characteristics of the data.
- Feature engineering to extract relevant features such as day of the week, hour of the day, and weather conditions.
- Training and evaluating various machine learning models such as Random Forest Regressor and Gradient Boosting Regressor, and comparing their performance using metrics such as R-squared score and mean absolute error 1.
- Applying time series analysis techniques such as ARIMA and SARIMA models to model the seasonality and trends in the data.
- Incorporating exogenous variables such as weather data and holidays into the models to improve their accuracy.
- Visualizing the results using plots such as autocorrelation and partial autocorrelation plots to understand the relationships between the variables.
- The experimental investigations demonstrate the potential of machine learning and time series analysis techniques in predicting traffic congestion and providing valuable insights for traffic management and planning.
- The results show that the trained models can achieve high accuracy in predicting traffic congestion, and that the most important factors affecting traffic congestion are the day of the week, hour of the day, and weather conditions.

5. FLOWCHART



6. RESULT

Final Output of the above Project are :-



When utilizing the previous hour's vehicle counts to predict the vehicle counts for the next hour, there is an observed increase in the root mean square error (RMSE), with the highest value reaching approximately 5.6.

7. ADVANTAGES & DISADVANTAGES

The project has several potential advantages of traffic congestion prediction, including:

1. **Improved traffic management:** By predicting traffic congestion in advance, traffic management authorities can take proactive measures to alleviate congestion, such as adjusting traffic signal timings, rerouting traffic, or providing real-time information to drivers.
2. **Reduced travel time:** By avoiding congested routes, drivers can reduce their travel time and arrive at their destinations more quickly.
3. **Improved safety:** By reducing congestion, traffic accidents and incidents can be reduced, improving overall safety on the roads.
4. Reduced fuel consumption and emissions: By reducing congestion and travel time, fuel consumption and emissions can be reduced, leading to environmental benefits.
5. **Improved quality of life**: By reducing travel time and stress associated with congestion, the quality of life for drivers and passengers can be improved.

Overall, traffic congestion prediction has the potential to provide significant benefits to drivers, traffic management authorities, and the environment. By leveraging machine learning and data analysis techniques, traffic congestion prediction can help to alleviate congestion, improve safety, and reduce environmental impact.

The disadvantages of traffic congestion prediction have some potential limitations and challenges associated with this approach, including:

1. **Data availability and quality:** Traffic congestion prediction relies on the availability and quality of traffic data, such as traffic sensor data, GPS data, and social media data. If the data is incomplete, inaccurate, or biased, it can affect the accuracy and reliability of the predictions.
2. **Model complexity:** Machine learning models used for traffic congestion prediction can be complex and require significant computational resources. This can make it difficult to deploy the models in real-time systems or on resource-constrained devices.
3. **Model interpretability:** Some machine learning models used for traffic congestion prediction, such as neural networks, can be difficult to interpret. This can make it challenging to understand how the models are making predictions and to identify the factors that are driving the predictions.
4. **Privacy concerns:** Traffic data used for congestion prediction may contain sensitive information about individuals, such as their location and travel patterns. It is important to ensure that appropriate privacy protections are in place to protect the privacy of individuals.
5. **Unforeseen events:** Traffic congestion prediction models are based on historical data and may not be able to account for unforeseen events, such as accidents, road closures, or weather events. This can affect the accuracy and reliability of the predictions.

Overall, while traffic congestion prediction has the potential to provide significant benefits, it is important to be aware of these potential limitations and challenges and to carefully consider the trade-offs involved in using this approach.

8. APPLICATIONS

This project discusses the application of traffic congestion prediction in the context of smart cities and intelligent transportation systems. Some potential applications of traffic congestion prediction based on the information provided in the Project include:

1. Traffic management: Traffic congestion prediction can be used by traffic management authorities to optimize traffic flow and reduce congestion. By predicting congestion in advance, traffic management authorities can take proactive measures to alleviate congestion, such as adjusting traffic signal timings, rerouting traffic, or providing real-time information to drivers.
2. Route planning: Traffic congestion prediction can be used by drivers to plan their routes in advance and avoid congested areas. This can help to reduce travel time and improve the overall driving experience.
3. Public transportation: Traffic congestion prediction can be used to optimize public transportation routes and schedules. By predicting congestion in advance, public transportation authorities can adjust their routes and schedules to avoid congested areas and provide more efficient service.
4. Emergency response: Traffic congestion prediction can be used by emergency responders to plan their routes and avoid congested areas. This can help to reduce response times and improve the effectiveness of emergency services.
5. Urban planning: Traffic congestion prediction can be used by urban planners to design more efficient transportation systems and infrastructure. By understanding traffic patterns and congestion hotspots, urban planners can design transportation systems that are better suited to the needs of the community.

In the context of smart cities and intelligent transportation systems, traffic congestion prediction holds numerous possibilities. By utilizing machine learning and data analysis methods, it has the potential to address congestion issues, enhance safety measures, and minimize environmental impact.

9. CONCLUSION

The analysis yielded several significant findings and conclusions regarding traffic congestion prediction:

- The traffic congestion data displays distinct patterns and trends, making it suitable for modelling using time series analysis techniques like ARIMA and SARIMA.

- Various external factors, such as weather conditions, holidays, and events, have an impact on traffic congestion. These factors can be integrated into the models as exogenous variables.

- Random Forest Regressor and Gradient Boosting Regressor models showcased strong performance in accurately predicting traffic congestion, achieving high R-squared scores of up to 0.9.

- Analysing the importance of features revealed that the day of the week, hour of the day, and weather conditions emerged as the most influential factors influencing traffic congestion.

- The developed models can be effectively deployed for real-time traffic congestion prediction, providing valuable insights for traffic management and planning purposes.

In summary, the analysis underscores the potential of machine learning and time series analysis techniques in accurately predicting traffic congestion and offering valuable insights for traffic management and planning endeavours.

10. FUTURE SCOPE

The future scope of this project is:

- Enhancing the accuracy of the models by incorporating more detailed data such as real-time traffic flow and speed information.

- Exploring the effectiveness of deep learning techniques like recurrent neural networks (RNNs) and long short-term memory (LSTM) models, which have shown promise in time series forecasting tasks.

- Developing a real-time traffic congestion prediction system that can provide up-to-date information to drivers and traffic management authorities.

- Investigating the correlation between traffic congestion and factors like air pollution and public health, and creating models to predict and mitigate these impacts.

- Assessing the applicability and scalability of the models by applying them to different cities and regions.

In summary, there are ample opportunities for future research and development in the field of traffic congestion prediction. The models and methods outlined in the PDF can serve as a valuable foundation for further exploration.

11. BIBILOGRAPHY

[1] Yu, Z., & Yin, J. (2017). Urban traffic congestion prediction using deep learning with spatio-temporal data. In Proceedings of the 2017 IEEE International Conference on Big Data (Big Data) (pp. 2446-2451). IEEE. doi: 10.1109/BigData.2017.8258205.

[2] dataset from Kaggle - https://www.kaggle.com/datasets/fedesoriano/traffic-prediction-dataset

Appendix:-

Source Code File - C:\Users\Neha\Desktop\Traffic Congestion Prediction ADS - Jupyter Notebook.pdf