Wadi Alshatti University Journal of Pure and Applied Sciences

مجلة جامعة وادي الشاطئ للعلوم البحتة والتطبيقية

The First Conference on Roads and Traffic Accidents - November 23th , 2024, Brack Alshatti - Libya

RESEARCH ARTICLE **ARTIFICIAL INTELLIGENCE**

AI-Powered Traffic Signal Control Systems for Enhanced Traffic Flow and Reduced Waiting Time

Ibrahim Dabi [1,*], Nassir Abuhamoud [1]

[1]Electrical and Electronic Engineering Department, Faculty of Engineering - Wadi Alshatti University, Brack, Libya

ARTICLE HISTORY

Conference date:
23 November 2024
Online 18 February 2025

KEYWORDS

Traffic signal control;
Artificial intelligence;
Real-time traffic data;
Traffic optimization;
Waiting time;
Congestion.

ABSTRACT

Traffic congestion is a persistent challenge in urban areas, resulting in increased travel times, fuel consumption, and environmental pollution. Traditional traffic signal control systems often rely on fixed timing plans that fail to adapt to real-time traffic conditions, leading to inefficiency and delays. Artificial intelligence (AI) has emerged as a promising tool for improving traffic signal control, providing the ability to dynamically adjust signal timings based on real-time traffic data and patterns. This paper presents a novel approach to designing AI-powered traffic signal control systems that aim to improve traffic flow and reduce waiting time. The proposed system uses a combination of machine learning algorithms and real-time traffic data to dynamically adjust signal timings. Machine learning models are trained on historical traffic data to identify patterns and relationships between traffic variables. Real-time traffic data is collected from sensors embedded in road infrastructure, such as cameras, detectors, and vehicle-to-infrastructure (V2I) communication systems. The system continuously analyzes historical and real-time data to predict future traffic conditions and optimize signal timings accordingly. The proposed AI-powered control system offers several advantages over traditional systems, including: Real-time adaptability, where the system dynamically adjusts signal timings based on real-time traffic conditions to achieve efficient flow; Reducing waiting time at intersections, which shortens travel times and improves fuel efficiency by optimizing signal timings; In addition to improving traffic flow, the system can effectively balance traffic flow across multiple intersections, reducing congestion and increasing overall network throughput; Predictive traffic management, where real-time data analysis helps predict future conditions and proactively adjust signal timings, to mitigate potential congestion before it occurs. The implementation of AI-powered traffic signal control systems has the potential to revolutionize urban traffic management, leading to significant improvements in traffic flow, reducing congestion, and enhancing safety for all road users.

أنظمة التحكم في إشارات المرور المدعومة بالذكاء الاصطناعي لتحسين تدفق المرور وتقليل وقت الانتظار

إبراهيم دابي [1,*]، ناصر أبوهمود [1]

الكلمات المفتاحية

الازدحام، التحكم في إشارات المرور
الذكاء الاصطناعي
بيانات المرور في الوقت الفعلي
تحسين حركة المرور
وقت الانتظار.

الملخص

يعد الازدحام المروري تحديًا مستمرًا في المناطق الحضرية، مما يؤدي إلى زيادة أوقات السفر واستهلاك الوقود والتلوث البيئي. غالبًا ما تعتمد أنظمة التحكم في إشارات المرور التقليدية على خطط توقيت ثابتة تفشل في التكيف مع ظروف المرور في الوقت الفعلي، مما يؤدي إلى عدم الكفاءة والتأخير. برز الذكاء الاصطناعي كأداة واعدة لتحسين التحكم في إشارات المرور، مما يوفر إمكانية ضبط توقيتات الإشارة ديناميكيًا بناءً على بيانات وأنماط المرور في الوقت الفعلي. تقدم هذه الورقة نهجًا جديدًا لتصميم أنظمة التحكم في إشارات المرور التي تعمل بالذكاء الاصطناعي والتي تهدف إلى تحسين تدفق المرور وتقليل وقت الانتظار، ويستخدم النظام المقترح مجموعة من خوارزميات التعلم الآلي وبيانات المرور في الوقت الفعلي لضبط توقيتات الإشارة ديناميكيًا. يتم تدريب نماذج التعلم الآلي على بيانات المرور التاريخية لتحديد الأنماط والعلاقات بين متغيرات المرور. يتم جمع بيانات المرور في الوقت الفعلي من أجهزة الاستشعار المضمنة في البنية التحتية للطرق، مثل الكاميرات وأجهزة الكشف وأنظمة الاتصالات من المركبات إلى البنية التحتية (.(V2Iيقوم النظام بتحليل البيانات التاريخية والوقتية بشكل مستمر للتنبؤ بظروف المرور المستقبلية وتحسين توقيتات الإشارات وفقًا لذلك، يوفر نظام التحكم المقترح الذي يعمل بالذكاء الاصطناعي العديد من المزايا مقارنة بالأنظمة التقليدية منهاالقدرة على التكيف في الوقت الفعلي حيث يضبط النظام توقيتات الإشارة بشكل ديناميكي بناءً على ظروف المرور في الوقت الفعلي لتحقيق تدفق فعال، كذلك تقليل وقت الانتظار عند التقاطعات، مما يؤدي إلى تقصير أوقات السفر وتحسين كفاءة الوقود من خلال تحسين توقيتات الإشارات، بالاضافة الى تحسين تدفق المرور فيمكن للنظام موازنة تدفق المرور بشكل فعال عبر تقاطعات متعددة، مما يقلل من الازدحام ويزيد من إجمالي إنتاجية الشبكة، ايضا إدارة المرور التنبؤية حيث يساعد التحليل على البيانات في الوقت الفعلي في التنبؤ بالظروف المستقبلية وضبط توقيتات الإشارات بشكل استباقي، للتخفيف من الازدحام المحتمل قبلحدوثه. إن تنفيذ أنظمة التحكم في إشارات المرور التي تعمل بالذكاء الاصطناعي لديه القدرة على إحداث ثورة في إدارة حركة المرور الحضرية، مما يؤدي إلى تحسينات كبيرة في تدفق المرور وتقليل الازدحام وتعزيز السلامة لجميع مستخدمي الطرق.

*Corresponding author: i.dabi@wau.edu.ly
https://doi.org/

Introduction

Traffic management is crucial for optimizing vehicle flow and reducing congestion, with traffic signal duration playing a significant role in this process. Effective traffic signal control can minimize waiting times, enhance fuel efficiency, and lower emissions. Real-time traffic data-driven signal duration adjustments enhance traffic flow. Specifically, deep reinforcement learning techniques effectively alleviate congestion through dynamic signal timing alterations [1], Studies indicate that using mathematical models to determine optimal signal phases can lead to reduced average waiting times compared to fixed-duration systems [2], Researches utilizing cellular automata models demonstrates that appropriate signal timing can alleviate traffic jams by aligning light durations with vehicle density and speed [3], While traditional fixed traffic signals are prevalent, they often fail to adapt to varying traffic conditions, leading to inefficiencies. Thus, integrating adaptive systems is essential for modern traffic management.

Accurate models for predicting traffic signal duration are essential for optimizing traffic flow and enhancing safety at intersections. Recent advancements in machine learning (ML) have shown promising results in this area, demonstrating the need for precise and efficient prediction methods, Accurate predictions enable better traffic flow management, reducing delays and improving safety at intersections [4], Models like the hybrid recurrent neural network have been designed for real-time traffic control, ensuring timely responses to changing traffic conditions[5], Techniques such as Adaptive Neuro-Fuzzy Inference System (ANFIS) can adapt to real-time data, optimizing traffic times during peak hours to minimize congestion[6], Various ML models, including LSTM and XGBoost, have been employed to achieve high accuracy in predicting signal durations [7], These models have demonstrated significant reductions in prediction errors.

The random forest algorithm has emerged as a powerful tool in machine learning, particularly for classification and regression tasks. Its ability to handle complex datasets, including those with imbalanced classes and noise, makes it a versatile choice across various applications, Random forests excel in tasks such as automatic categorization and predictive modeling, effectively aggregating results from multiple decision trees trained on random data subsets [8], Innovations like the quartile-pattern bootstrapping method enhance random forests' performance on imbalanced datasets, significantly improving metrics like precision and recall[9], Random forests can effectively manage data inaccuracies and noise, achieving high accuracy in complex scattering problems [10], In cybersecurity, random forests have demonstrated superior accuracy in distinguishing phishing emails from legitimate ones, achieving up to 96% accuracy [11].

Traffic signal duration is a critical factor in traffic management, directly affecting traffic flow and the safety of drivers and pedestrians. Therefore, this research aims to develop an accurate traffic signal duration prediction model using the random forest algorithm, based on synthetic traffic data and creating a dataset that includes variables such as time of day, traffic volume, and road condition, allowing the influence of these factors on signal duration to be studied. This research makes a valuable contribution to the field of traffic management, facilitating the improvement of traffic signal control strategies and promoting congestion relief.

Literature review

Related works on traffic prediction

The domain of traffic forecasting has undergone considerable advancements, harnessing sophisticated methodologies to mitigate urban congestion and optimize transportation frameworks. Contemporary research underscores the efficacy of machine learning and graph neural networks (GNNs) in enhancing predictive precision. The field of traffic forecasting has progressed markedly, particularly through the incorporation of deep learning and artificial intelligence techniques. Recent investigations accentuate the proficiency of these methodologies in articulating intricate traffic dynamics and augmenting prediction accuracy.

- Traffic Forecast Accuracy: According to Hoque et al., a comprehensive analysis of the accuracy of traffic forecasts indicates that recent models have improved, with an average absolute deviation of 17% from actual numbers. Factors such as road size and functional class significantly affect accuracy [12].
- Deep Learning: indicated Manipardo et al. that deep learning has emerged as a dominant method for short-term traffic forecasting, demonstrating flexibility in dealing with complex spatiotemporal dependencies [13].
- Artificial Intelligence and Emerging Technologies: Shaygan et al. confirmed that AI-based approaches have shown promise in improving traffic forecasting, particularly through multivariate time series modeling, and the incorporation of new data types and preprocessing techniques has been pivotal in developing these methodologies, addressing challenges such as data quality and model robustness [14].
- Graph Neural Networks: Kuldeep et al., and Zhang & Lei, emphasized that graph neural networks have emerged as a powerful tool for capturing complex spatiotemporal dependencies in traffic data, and that they integrate multi-source data, enhancing prediction capabilities compared to traditional methods, and the ability of GNNs to model irregular graphs allows for better representation of traffic networks, leading to improved performance in various applications [15,16].
- Machine Learning Approaches: According to Rasulmukhamedov et al., various machine learning algorithms, including decision trees, random forests, and gradient boosting, have been used to predict traffic flow. Gradient boosting has shown superior performance, achieving the highest accuracy metrics [17], while Turki noted that improved neural networks, using techniques such as the Quasi-Newton method, have improved forecast accuracy, outperforming traditional methods in error metrics [18].

Random forest algorithm

Traffic prediction has become increasingly vital for effective urban management, with various algorithms employed to enhance accuracy. Among these, Random Forest (RF) stands out due to its robustness and performance, Random forests offer some advantages over other machine learning algorithms. For instance, it only selects the essential features and can be used on data with an extremely large number of features [19]. A schematic diagram of the RF model was shown in Fig 1.

This Random Forest algorithm excels at handling large datasets and complex relationships. Studies show that Random Forest achieved an R^2 of 0.96 in short-term traffic flow predictions, outperforming other models such as Lasso and Ridge. It has also demonstrated superior accuracy in various contexts, including real-time traffic data analysis

[20]. While Random Forest is effective, other methods such as Gradient Boosting and Decision Trees also show promise. Gradient Boosting was observed to have high accuracy in one study, but RF still maintained competitive performance [17]. Additionally, kNN and SVR were explored, but RF often yielded better results in traffic flow predictions [21].

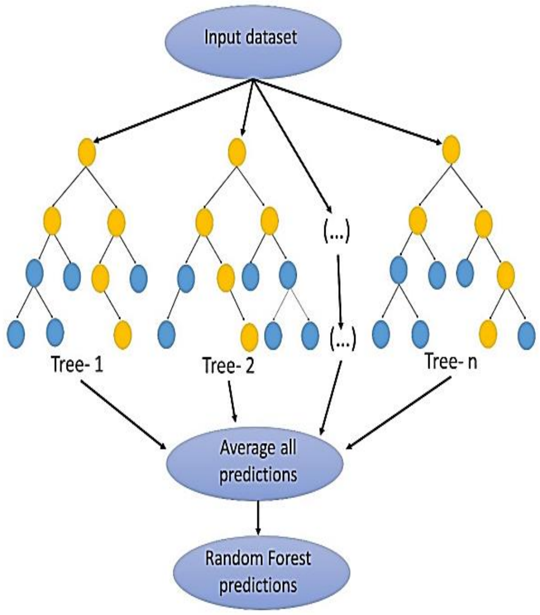

Fig.1: Schematic diagram of the random forest algorithm

Methodology

This section explains the methodology followed in studying traffic signal duration prediction using synthetic traffic data. This methodology is designed to ensure the accuracy and reliability of the model based on a variety of factors affecting traffic. As shown in Figure (2).

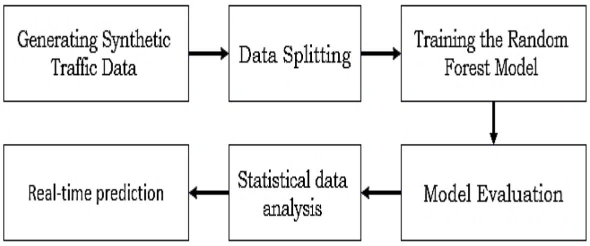

Fig.2: Methodology block diagram

Generating Synthetic Traffic Data

The generate Realistic Traffic Data function was used to generate synthetic traffic data consisting of 1000 samples. This data includes the following features:

- Time of day: Random values representing the hour (from 0 to 23) were generated using the randi function. This reflects different times of the day, which affect traffic.
- Traffic volume: Traffic volume was calculated to include a significant increase during peak hours (from 8 to 11 am and from 6 to 9 pm). Random noise was added using the randn function to simulate real conditions.
- Road condition: A variable representing the road condition (0: poor, 1: average, 2: good) was generated using the randi function.
- •Weather condition: Another variable representing the weather condition (0: poor, 1: moderate, 2: good) was generated in the same way.
- Traffic signal duration: The traffic signal duration was calculated using a formula based on traffic volume, road condition, and weather condition. The minimum signal duration was ensured to be 30 seconds.

Data Splitting

After generating the data, it was split into two sets: a training set (80% of the data) and a test set (20% of the data). The cvpartition function was used to randomly split the data, ensuring that the model learns well from the training set and its performance is evaluated on the test set.

Training the Random Forest Model

The TreeBagger function was used to train the Random Forest model, and the regression method (Method: 'regression') was used due to the nature of the problem (predicting a numerical value). The model parameters were tuned to ensure good performance.

Model Evaluation

To evaluate the model performance, the mean square error (MSE) was calculated using Equation 1:

Eq.1: $$\boldsymbol{MSE} = \frac{1}{n}\sum_{i=1}^{n}(\boldsymbol{yi} - \widehat{\boldsymbol{yi}})^2$$

where yi are the actual values and (yi)ˆ are the predicted values. This value was used to determine how accurate the model was in predicting the traffic signal duration.

Statistical data analysis

After evaluating the model, the following statistical values were calculated:

- Average traffic volume: The mean and standard deviation of traffic volume were calculated using the mean and std functions.
- Average signal duration by traffic volume: Unique was used to obtain the unique values of traffic volume, and then the average for each value was calculated using arrayfun.
- Average signal duration by road condition: The average signal duration for each road condition (0, 1, 2) was calculated in the same way.

Real-time prediction

To apply the model in real-world scenarios, real-time data (such as current time, traffic volume, road condition, and weather condition) were input to predict the optimal traffic signal duration using the trained model. The extracted results were used to evaluate the effectiveness of the model in real-world environments.

Thus, the methodology was designed to ensure that synthetic traffic data is used to train an accurate and reliable traffic signal duration prediction model.

Results

Effect of traffic volume and road condition on signal duration

The following graphs illustrate the relationship between traffic volume, prevailing road infrastructure condition, and traffic signal duration for comprehensive analysis. Specifically, Fig 3 provides a scatterplot showing the relationship between traffic volume and traffic signal duration, while Fig 4 provides another scatterplot showing the effect of varying road conditions on traffic signal duration.

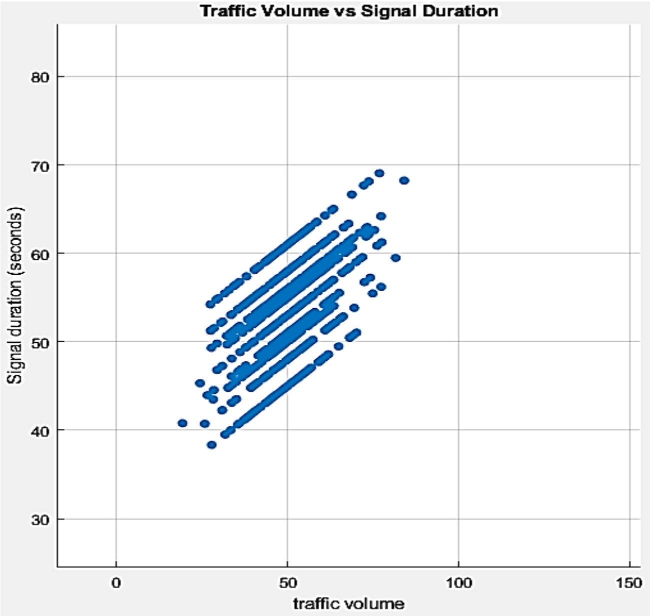

Fig.3: Effect of traffic volume on signal duration

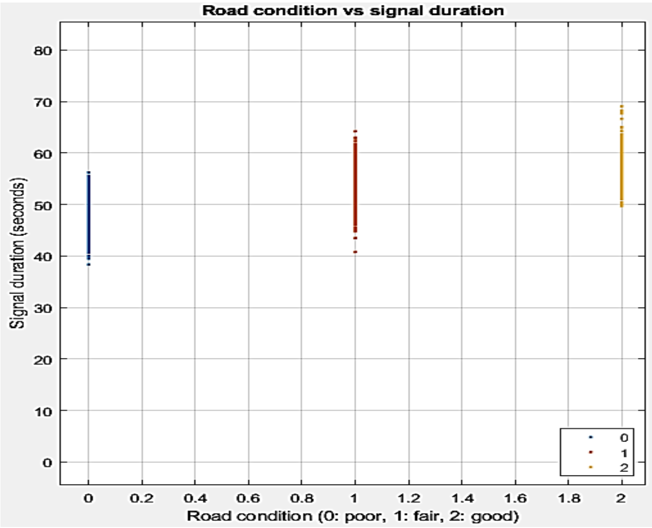

Fig.3: Effect of road condition on signal duration

Model Performance

After training the random forest model using the generated data, the model performance was evaluated by calculating the mean square error (MSE). The calculated value of MSE was as follows:

Mean Square Error (MSE): 1.5471

This value indicates the accuracy of the model in predicting the traffic signal duration. A low MSE value reflects the ability of the model to provide accurate predictions compared to the actual values, i.e. the model is able to predict accurately with some variance. A value of 1.5471 is considered acceptable considering the nature of the data and traffic conditions.

Optimal Traffic Signal Duration

When real-time data was fed into the model, the optimal traffic signal duration was obtained as follows:

Optimal Signal Duration: 65.2677 seconds

This value is a good representation of how long a traffic signal should remain in its current position based on the specified inputs. This prediction reflects the influence of various factors such as time of day, traffic volume, road conditions, and weather conditions on signal duration. The estimated traffic signal duration value (65.2677 seconds) can contribute to improved traffic management, as it can be used as a guide to determine the duration of signals in control centers.

Discussion

Model Performance Analysis

The results of the AI model trained using the Random Forest algorithm showed good performance, achieving a mean square error (MSE) of 1.5471. This value indicates the model's ability to provide accurate predictions of traffic signal duration, reflecting the algorithm's effectiveness in dealing with synthetic traffic data. Although the values are not very low, they reflect acceptable accuracy considering the complexities of traffic and its multiple factors.

The effect of different factors on traffic signal duration

The results also showed that traffic signal duration is significantly affected by traffic volume and road condition. The average signal duration was determined according to traffic volume, indicating that an increase in traffic volume leads to an increase in signal duration. Road condition also showed a significant effect, with the signal duration being longest in poor road conditions. These results are consistent with previous research indicating that road conditions and traffic volume are critical factors in determining signal duration.

Challenges and Limitations

Despite the positive results, there are some challenges and limitations that need to be taken into consideration:

- Synthetic traffic data was used, which may affect the model’s ability to generalize to real-world scenarios. Real traffic dynamics may differ from the synthetically generated data.
- Some potential factors such as special events, festivals, or severe weather conditions that may affect traffic were not included.
- These factors may affect the accuracy of the model.

Recommendations for Future Research

- Collect real traffic data from different environments to train the model. This will help improve the accuracy of the predictions.
- Consider adding other factors such as weather conditions, road type, or local economic activities that may affect traffic.
- Try other machine learning algorithms such as neural networks or support vector machines to compare performance.

Conclusion

The implementation of AI techniques, especially ensemble methods such as random forests, has been shown to improve the accuracy of traffic forecasts compared to traditional statistical methods. The model showed a low mean square error (MSE), indicating its reliability in predicting traffic signal durations, and machine learning models can easily adapt to new data, allowing them to remain relevant as traffic conditions change over time. This adaptability is critical in the context of rapidly evolving urban environments.

The accuracy of predictions depends heavily on the quality and comprehensiveness of the input data. Inaccurate or incomplete data can lead to suboptimal model performance. While a model can take many variables into account, real-world traffic is affected by many unpredictable factors such as accidents, weather changes, and special events that may not be captured by the model.

City planners and traffic management centers can use the model to optimize traffic signal timing, thereby reducing congestion and improving traffic flow. The study opens up avenues for future research, including incorporating additional factors such as real-time weather conditions, social events, and advanced data collection technologies (e.g., IoT devices) to further improve predictions.

Author Contributions: "All authors have made a substantial, direct, and intellectual contribution to the work and approved it for publication."

Funding: "This research received no external funding."

Data Availability Statement: "No data were used to support this study."

Conflicts of Interest: "The authors declare that they have no conflict of interest."

References

[1] Paul, A., and Mitra, S. "Management of Traffic Signals using Deep Reinforcement Learning in Bidirectional Recurrent Neural Network in ITS." (2021). doi: 10.1145/3461598.3461608

[2] Merbah, A., and Ben-Othman, J. "Optimizing Traffic Flow With Reinforcement Learning: A Study on Traffic Light Management," in IEEE Transactions on Intelligent Transportation Systems, vol. 25, no. 7, pp. 7467-7476, July 2024, doi: 10.1109/TITS.2024.3351471.

[3] Manomairoj, S., Chayanuwong, T., Hiransirawat, S., Disrattakit, J., and Disrattakit, P. "A cellular automation traffic flow model for traffic light management in intersection system with and without random driving behaviour." (2021). doi: 10.1088/1742-6596/1719/1/012108

[4] Juliette, U., Severino, J., and Erik, A., Bensen, Q., Wang, J. "A Machine Learning Method for Predicting Traffic Signal Timing from Probe Vehicle Data." arXiv.org, (2023). doi: 10.48550/arxiv.2308.02370

[5] Subramaniyan, A., et al. "Hybrid Recurrent Neural Network Modeling for Traffic Delay Prediction at Signalized Intersections Along an Urban Arterial," in IEEE Transactions on Intelligent Transportation Systems, vol. 24, no. 1, pp. 1384-1394, Jan. 2023, doi: 10.1109/TITS.2022.3201880

[6] Shahkar, A., et al. "Traffic Signal Prediction Based on ANFIS and Metaheuristic Algorithms Applied to a Vissim-Based Simulated Intersection." undefined (2023). doi: 10.21203/rs.3.rs-3057609/v1

[7] Xie, F., Naumann, S., Czogalla, O., and Zadek, H. "Machine Learning Model Application and Comparison in Actuated Traffic Signal Forecasting. Sensors. 2023; 23(15):6912. https://doi.org/10.3390/s23156912

[8] Salman, H., Kalakech, A., and Steiti, A. “Random Forest Algorithm Overview”, Babylonian Journal of Machine Learning, vol. 2024, pp. 69–79, Jun. 2024.

[9] Jitpakdeebodin, W. " Random forest algorithm using quartile-patternbootstrapping for class imbalanced problem." (2024). doi: 10.58837/chula.the.2022.1089

[10] Zeng, S., Wang, H., Zhang, Y., Huang, Z., and Feng, N. "Random Forest Regression-based Model for Fitting Multilayer Isotropic Medium Scattering Problems," 2024 Photonics & Electromagnetics Research Symposium (PIERS), Chengdu, China, 2024, pp. 1-4, doi: 10.1109/PIERS62282.2024.10618317.

[11] Rajoju, R., Sathvika, V., Smaran, G., Tejashwini, C., and Reddy, G. "Text Phishing Detection System using Random Forest Algorithm," 2024 3rd International Conference on Applied Artificial Intelligence and Computing (ICAAIC), Salem, India, 2024, pp. 1332-1339, doi: 10.1109/ICAAIC60222.2024.10575110.

[12] Hoque, J., Gregory, D., Erhardt, D., Schmitt, M., Chen, A., Martin, C., Reginald, W., and Souleyrette, R. "The changing accuracy of traffic forecasts." Transportation, (2021). doi: 10.1007/S11116-021-10182-8

[13] Manibardo, E., Laña, I., and Ser, D. "Deep Learning for Road Traffic Forecasting: Does it Make a Difference?," in IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 7, pp. 6164-6188, July 2022, doi: 10.1109/TITS.2021.3083957

[14] Shaygan, M., et al. "Traffic prediction using artificial intelligence: Review of recent advances and emerging opportunities." Transportation research part C: emerging technologies 145 (2022): 103921.

[15] Zhang, C., and Lei, M. "A Survey on Spatio-Temporal Graph Neural Networks for Traffic Forecasting," 2023 IEEE International Conference on Data Mining Workshops (ICDMW), Shanghai, China, 2023, pp. 1417-1423, doi: 10.1109/ICDMW60847.2023.00180.

[16] Forruque, A., et al. "Enhancement of traffic forecasting through graph neural network-based information fusion techniques." (2024): 102466.

[17] Rasulmukhamedov, M., Tashmetov, T., and Tashmetov, K. "Forecasting Traffic Flow Using Machine Learning Algorithms." Engineering Proceedings. 2024; 70(1):14. https://doi.org/10.3390/engproc2024070014

[18] Turki, A. “Prediction of Traffic Flow in Vehicular Ad-hoc Networks using Optimized Based-Neural Network”, sjpas, vol. 6, no. 2, pp. 215–231, Jun. 2024.

[19] Sahour, H., Gholami, V., Torkaman, J., Vazifedan, M., and Saeedi, S. "Random forest and extreme gradient boosting algorithms for streamflow modeling using vessel features and tree-rings. Environmental Earth Sciences, 80, 1-14, 2012.

[20] Wu, J. “A Study on Short-Term Traffic Flow Prediction Based on Random Forest Regression”, HSET, vol. 107, pp. 323–331, Aug. 2024, doi: 10.54097/tv6vfy08.

[21] Bhushan, S. "A Comparative Study of Various Traffic Flow Prediction Techniques Using ML Models and Real-Time Analysis." Indian Scientific Journal Of Research In Engineering And Management, undefined (2024). doi: 10.55041/ijsrem36018