

An XGBoost-SHAP Framework for Predicting Traffic Speed Based on Road Geometry



O. Bayode^{1*}, O. Aiyelokun^{1,2}, A. A. Olanipekun¹, D.I. Idusuyi¹, E.O. Akande¹

¹Department of Civil and Environmental Engineering, Bells University of technology, Ota, Ogun State, Nigeria

^{1,2}Department of Civil Engineering, University of Ibadan, Oyo State, Nigeria



ABSTRACT: In Intelligent Transportation System (ITS), Road Traffic Speed Prediction (RTSP) is essential for effective road operation and safety. This study investigates the impact of highway geometric features on traffic operating speed of Idiroko Road, a critical transportation link in Nigeria with the border. Geometric features including radius of curvature, road gradient, pavement width, shoulder width, median width and the presence of intersections were determined on site, as well as the spot speed of different vehicle types using a stopwatch approach. The Extreme Gradient Boosting (XGBoost) model was used to predict traffic operating speed, and Shapley additive explanations (SHAP) were employed to interpret the correlation between variables. The findings reveal that the radius of curvature, road gradient, pavement width, shoulder widths and the presence of intersections along the road have a significant effect on traffic operating speeds. The XGBoost model demonstrated high accuracy and efficiency with 98.5% and 97.9% coefficient of determination (R^2) in both testing and training datasets. This study emphasizes the effectiveness of XGBoost in predicting traffic operating speed. The findings can be implemented in urban road design, infrastructural planning and traffic management ultimately contributing to safer and more efficient road networks.

KEYWORDS: Highway Geometric, Traffic Operating Speed, Speed Prediction, XGBoost Model, SHAP

[Received Mar. 20, 2025; Revised June 1, 2025; Accepted June 9, 2025]

Print ISSN: 0189-9546 | Online ISSN: 2437-2110

I. INTRODUCTION

With the increase in urbanization and motorization, road traffic demand continues to rise significantly. Intelligent Transportation Systems (ITS) have enhanced road functionality and efficiency by employing machine-learning models to predict traffic flow, speed, and safety. Traffic speed prediction is a crucial aspect of transportation engineering, influencing road safety, traffic management, and infrastructure planning. Accurate speed prediction models enhance traffic efficiency and reduce accident risks (Ganan-Cardenas *et al.*, 2025). Highway geometries including radius of curvature, relationship highlights the necessary consideration in highway design, planning and maintenance (Fitzpatrick, 2015; Bayode *et al.*, 2024).

Traditional statistical methods such as linear, multiple linear, Poisson, and binomial regression models have been employed by transportation researchers for traffic speed prediction (Semeida, 2013; Charlton *et al.*, 2014; Donnell, 2016; Montella *et al.*, 2020). Charlton *et al.* (2014) examined driver behavior in relation to speed, focusing on the influence of perceived risk and comfort. Their findings indicate that variations in speed choices are significantly influenced by the geometric design of the roadway. Donnell (2016) investigated the link between road curvatures, traffic speed and accident and reflect underlying spatial factors in traffic state prediction (Lee *et al.*, 2021; Abdelraouf *et al.*, 2022; Min *et al.*, 2023). Imoh and Rad (2025) analyzed and predicted traffic condition

gradient of road, pavement width, shoulder width and road intersection are essential for the regulation of traffic flow, traffic speed, road safety, and the behaviour of motorists (drivers navigate highways and decision-making), thereby affecting their perception of safety, visibility, and comfort (Semeida, 2013; Bouktif *et al.*, 2020). The complexity and multidimensional relationship between road geometric design including horizontal and vertical alignments, sight distance, changes in lane and shoulder widths and driver behaviour affects drivers exhaustion, speed choices and reaction time when driving (Dong *et al.*, 2015; Yang *et al.*, 2022), thereby affecting the overall safety of traffic along the road. This

rates along two-lane highway using linear regression model. The study observed increase in accident frequency at sharper curves and with high traffic speed. In addition, Montella *et al.*, (2020) found that gradient and horizontal alignment significantly influence speed consistency and safety outcomes. However, recent transportation researchers employ machine-learning models for the prediction of traffic operations on highway and compare the performance with the traditional models (Yang *et al.*, 2022; Ma *et al.*, 2023; Lee, 2023; Imoh and Rad, 2025). Convolutional neural network (CNN), Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) models have been employed to forecast travel time and traffic speed, accounting for spatio-temporal characteristics on road network in Ikorodu, Lagos state, Nigeria using some machine learning models including decision trees, gradient boosting and random forest classifier. The study classified the

*Corresponding author: ope.olowo.oo@gmail.com

doi: <http://dx.doi.org/10.63746/njtd.v22i3.3514>

traffic condition into peak and off-peak periods. Bayode *et al.* (2025) applied support vector machine (SVM) and extreme gradient boost (XGBoost) for road safety prediction on some road networks in SouthWest, Nigeria considering human factors, traffic operations and the road geometry. The result highlighted the factors contributing to road traffic accident along the route. Ma *et al.* (2023) also used Bayesian deep multi-linear relationship network to predict traffic flow and speed on selected road networks. These studies illustrates the effectiveness and efficiency of machine learning models as fast and reliable tools in traffic operation predictions.

Machine learning boosting models such as adaptive boosting (AdaBoost), extreme gradient boost (XGBoost) and Categorical boost (CatBoost) are likewise explored in the prediction of road traffic operations, accident injury and improve prediction accuracy (Parsa *et al.*, 2020; Ma *et al.*, 2021; Yang *et al.*, 2022; Lee, 2023). Extreme gradient Boosting model (XGBoost) in particular offer improved accuracy and robustness prediction with less underfitting and overfitting of data.

Lee, 2023 explored the use of XGBoost model and SHAP explanatory to predict road traffic operation along freeway considering the effect of intersection. Yang *et al.* (2022) utilized XGBoost in the prediction of road traffic accident severity considering several influencing factors. The study highlights road alignment and condition as a significant factor contributing to road accident severity. Parsa *et al.*, (2020) applied XGBoost and SHAP for real-time accident detection. The results shows that XGBoost can detect accidents robustness including accuracy, false alarm rate and accident occurrence.

However, limited research has been conducted on the use of machine learning boosting models to predict traffic operating speed, particularly with consideration to road geometry parameters. Due to the accuracy and efficiency of this boosting model, it is important to explore the performance in traffic speed prediction. This research explores the application of the XGBoost model in predicting traffic operating speeds while considering road geometry as a critical factor. The impact of highway geometric design on traffic operating speed is crucially illustrated using Idiroko Road, Ado- Odo, Ota, Ogun State, a significant highway in Nigeria. A critical transportation link, this road facilitates the movement

of products and individuals between Nigeria and its neighboring countries. However, vehicles encounter substantial obstacles due to the geometric characteristics of the road, which include a sharp curve, inconsistent shoulder provisions, and varying lane widths. This study investigated the relationship between the traffic operating speed and the geometric characteristics of the highway on Idiroko Road Ado-Odo, Ota, Ogun State, and predicts the traffic operating speed along the route. The ultimate goal is to improve traffic flow and enhance traffic safety.

II. METHODOLOGY

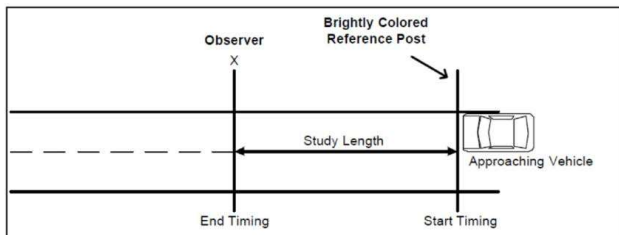
The study utilized an urban multilane highway, Idiroko road, Ado- Odo, Ota in Ogun State, Nigeria. The highway is major highway connecting the boarder of the country. Several industrial companies and vehicle types often use the road leading to traffic congestion. It connects 15 intersections and transit length of 200 km with deign speed of 100 km/hr. Figure 1shows the road network extracted from google earth map. The methodology adopted in this research to predict the operating speed prediction is divided into four main steps: (1) Data collection, (2) Data preprocessing, (3) XGBoost models, and (4) Performance Evaluation.

A. Data Collection

The data collected was categorized into speed data and road geometric data. 150 points were selected on the highway for the analysis. The spot speed which refers to the instantaneous speed of a vehicle at a specific location of different vehicle type (passenger car, 2axle-4tyres, 2axle-6tyres, 3axles and 4axles) were determined at each point using stopwatch approach (AASHTO, 2014) see figure 2. The number of vehicle collected at each point ranges from 20 to 50 accumulating to 7450 spot speed for all vehicle type. The road geometric characteristics considered in this study includes pavement width, shoulder width, median width, curve radius and road gradient were obtained on site (AASHTO, 2012). Table 1 shows the variable symbol and descriptive statistics of data. All data were obtain on site via site investigation, at daylight and good weather condition.



Figure 1: Map of study area extracted from Google Earth Map

Figure 2: Spot Speed Determination adapted from Roess *et al.* (2011)

B. Data Pre-processing

The performance of model depends on the relevance input parameters; therefore, irrelevance parameters introduce noise or bias to the model (Yang et al., 2022; Yin et al., 2023). The collected data undergoes preprocessing, data cleaning to removing duplicate entries, handling missing values, and correcting inconsistencies in the dataset. The datasets were splinted into training (70%) and testing (30%) sets.

Table 1: Descriptive Statistics of Independent Variables

Variables	Max	Min	Mean	S.D
PW	7.8	7.4	7.541	0.162
SW	2.71	2.1	2.478	0.239
NL	4	2	3	-
MW	0.8	0.44	0.46	4.92E(-16)
INT	1	0	-	-
R	750	1328	733	556
G	-6.5	10	0.54	2.8

PW (Pavement Width), SW (Shoulder Width), NL (Numbers of Lane), MW (Median Width), INT (Intersection), R (Radius of Curvature) and G (Gradient of Vertical Curve).

B. Extreme Gradient Model Implementation

The Extreme Gradient Boosting (XGBOOST) algorithm was implemented in this research to investigate the relationship between highway geometric properties and operating speed. The selection of Extreme boost was motivated by its capacity to generate precise and dependable predictions and its exceptional reliability in managing structured data. Lee (2023) and Yang *et al.* (2022) described XGBoost as powerful machine learning algorithm that improves prediction accuracy through a boosting framework. The fundamental of XGBoost as developed by Chen and Guestrin (2016), the model builds multiple decision trees sequentially, minimizing errors by adjusting weights and optimizing loss functions. The model predicts the target variable (\hat{y}_i) from ensemble decision trees by K additive function expressed in Equation 1. The Objective function of the model expressed in Equation 2 combines the loss function which measures the model fitness and the regularizer responsible to control the overfitting complexity of the model. The regularizer also includes two penalties; penalty of tree numbers and leaf weights. The values assigned to the terminal nodes of each tree are the leaf weight representing the predicted output of samples in the leaf. These penalties are determined using Taylor's second order approach in Equation 3 and gradient descent to optimise the objective function. Equation 4 shows the computed regularizer.

$$\hat{Y} = \Phi(x_i) = \sum_{k=1}^n f_k(x_i), f \quad (1)$$

$$Obj(\Phi) = \sum l(y_i, \hat{y}_i) + \sum \Omega(f_n) \quad (2)$$

$$f(x + \Delta x) = f(x) + f^1(x)\Delta x + \frac{1}{2}f^{11}(x)\Delta x^2 \quad (3)$$

$$Obj(\Phi) = \sum \{l(y_i, \hat{y}_i) + g_i f(x_i) + \frac{1}{2} h_i f^2(x_i)\} + \sum \Omega(f_n) \quad (4)$$

Where: \hat{y}_i is the target predicted variable, y_i is the independent variable of computed operating speed, x_i are the dependent variables selected geometric features, n are the numbers of features, f_k , is the tree structure with leaf score, f is the space of trees, Ω is the regularizer and g_i, h_i are the optimization value.

1) Hyperparameter Tuning Optimization and SHAP additive explanation

Hyperparameter tuning was employed in the study to optimize the XGBoost model using Grid Search. This was developed by analyzing dependencies between road geometry and speed predictions. Tuning of XGBoost prevents overfitting and minimize model complexity (Sun *et al.*, 2021). The optimization results are to show the models learning rate which are weight of each tree, learning step and process, the Max_depth illustrating the maximum depth of trees, n_estimator indicating the control numbers of weak estimators and the subsample, that is the random sampling ratio. In addition, SHAPly Additive Explanation (SHAP) refines these predictions by considering probabilistic dependencies between road geometry and speed (Lee, 2023). SHAP operated using Equation 5 This hybrid model improves interpretability and adaptability under varying conditions. The model outputs speed predictions based on road geometry ranking of each dependent variable on the impact on traffic operating speed output (Khosbayan *et al.*, 2021).

$$\phi_m = \frac{p^{!}(m-p/-1)!}{m!} [v(p\mu\{1\} - v(p))] \quad (5)$$

C. Model Performance Evaluation

The performance of the XGBN model was assessed for both the training and testing dataset using Mean Absolute Error (MAE) utilizing Equation 6 to evaluate the average error magnitude. Root Mean Squared Error (RMSE) measures the prediction accuracy utilizing Equations 7, Coefficient of determinant R-squared (R^2) to determine the goodness of fit Equations 8 and Correlation (COR) to analyzes prediction accuracy for different speed categories (Amiri *et al.*, 2020; Tayebikhorami and Sacchi, 2022).

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y} - y_i| \quad (6)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y} - y_i)^2} \quad (7)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y} - y_i)^2}{\sum_{i=1}^n (\hat{y} - \bar{y})^2} \quad (8)$$

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i \quad (9)$$

Where; n is the number of observations, y_i is the actual value of the i th observation, \hat{y} is the predicted value of the i th observed and \bar{y} is the average value of observations computed with Equation (9).

III. RESULTS AND DISCUSSION

A. Traffic Speed Cumulative of Vehicle type

The cumulative percentile speed indicates the percentage of vehicles using the road at different speed and operating speed represents the speed at which drivers are typically observed operating their vehicles (Berhanu *et al.*, 2024). Figure 2 shows the cumulative percentile of the four different vehicular types along the study route indicating the 50th speed percentile, which is the average traffic speed, the 85th speed percentile (V85) known as the operating speed and the 90th percentile. For passenger car, 2axle-4tyres, 2axle-6tyres, 3axles and 4axles

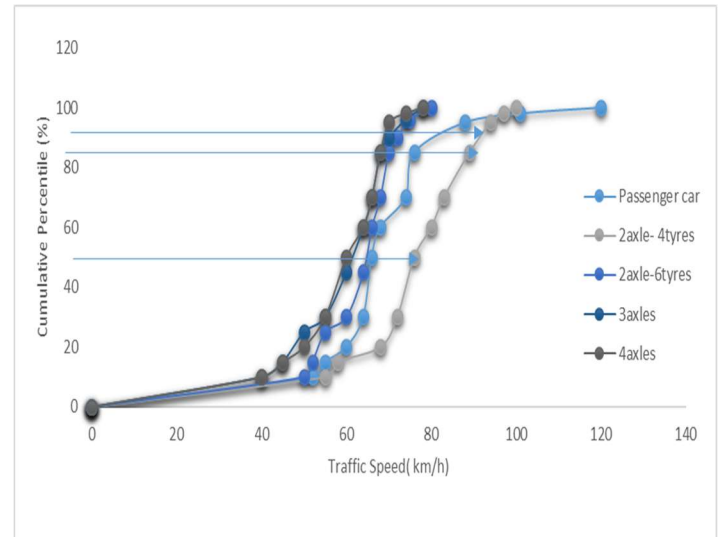


Figure 3: Cumulative Percentile Speed of Different Vehicle.

B. Variable Interaction and Interpretation

The SHAP values of seven features of the XGBoost model in figure 4(a) illustrates the interaction between variables. The variables were arranged according to their importance to the traffic speed prediction. The SHAP value implicit the impact of the variables on each other as the larger the SHAP value, the greater the impact and otherwise. The concentration of the dots denotes the strength of interaction, the red dots are mainly concentrated on the side with SHAP values less than 0 and the

blue dots are mainly concentrated on the side with SHAP values greater than 0 (Ahmed *et al.*, 2023; Chen and Chen, 2020). The interactions are not linear; R is negatively correlated with G with the correlation lightly displaying positive. G and NL have a weak positive correlation. Non-linear interaction exists between INT and R, also PW and SW. The interaction effects between PW, SW and MW are positive correlated.

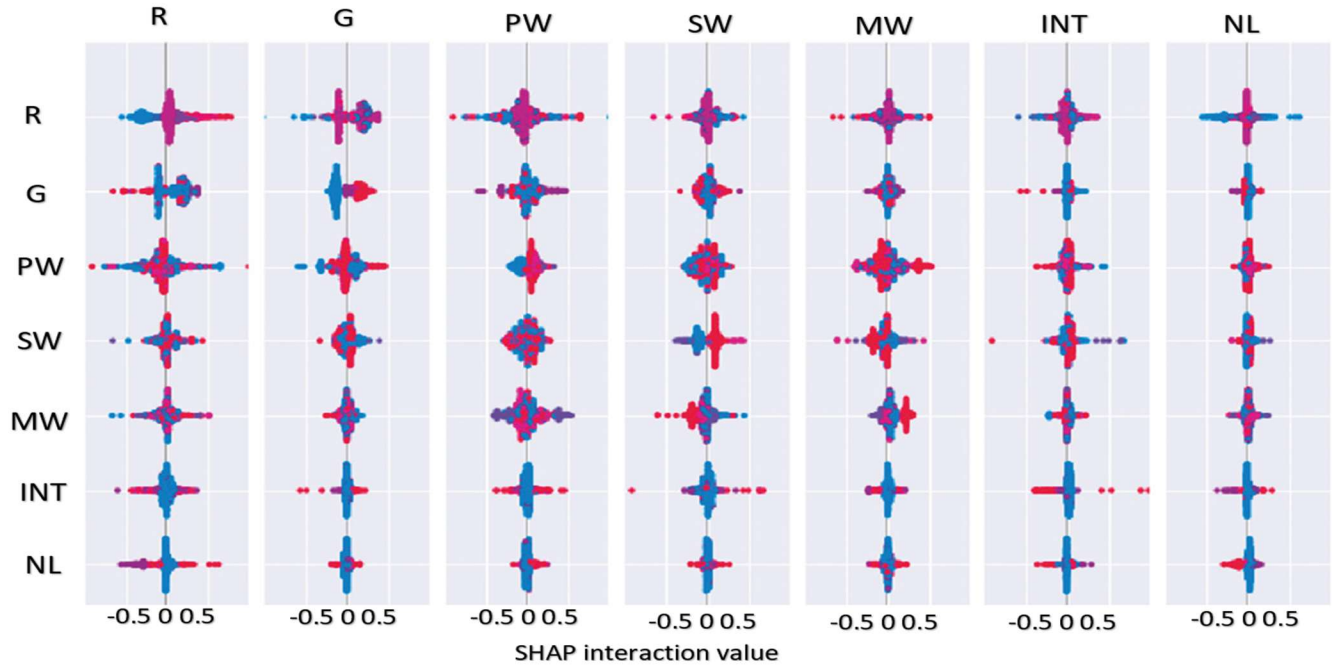


Figure 4(a): Interaction plot for Independence Variable

Figure 4(b) describes a heat map of correlation coefficients between variables to independent variable. The correlation coefficient shows the importance of variable to the prediction of traffic speed, all variables are sorted in order from positive to negative from radius of curvature, grade of road, numbers of lane, intersection, pavement width, shoulder width and median width. The Radius of curvature, presence of intersection, pavement width and shoulder width has significant positive correlation with traffic speed. This is reasonable because as vehicle approaches curve and intersection the speed tends to reduce (Shiran *et al.*, 2021; Mystakidis *et al.*, 2025). Sharper curves reduce operating speed due to increased lateral friction and driver caution. While the presence of intersections lowers

operating speed due to potential conflicts and required stopping behavior. Wider lanes facilitate higher speeds, reducing vehicle interaction, that is, the wider the width of the road and the shoulder the higher the vehicular speed. The steeper the grade of the road leads to decrease in speeds for heavy vehicles, whereas declines may increase speeds, potentially raising safety concerns. The median width, numbers of lanes has less significant in traffic speed prediction. The practical implications of road geometry influence on traffic speed includes urban road design, infrastructural planning and traffic management. The model optimizes speed limits based on geometric design, infrastructural design and road authorities to improve traffic operation and safety.

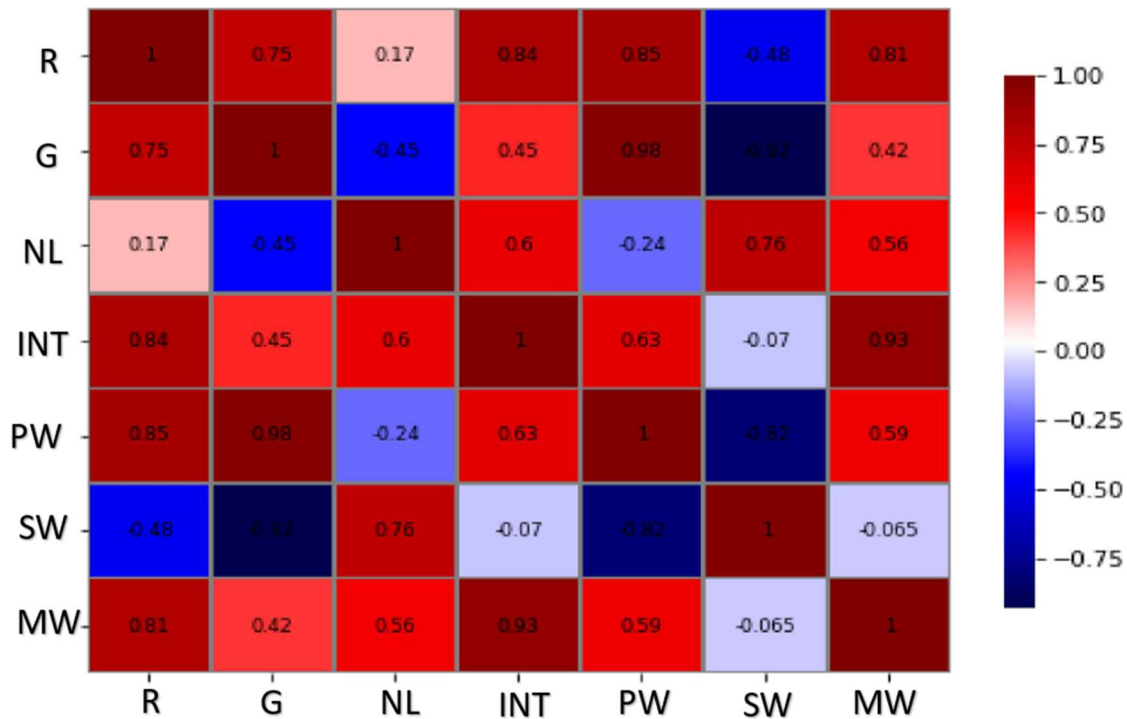


Figure 4(b): Heat Map of Variables

C. Model Performance Evaluation

Table 2 shows the summary of the result of the performance evaluation conducted on the model using R2, MAE, MSE, RMSE and COR criteria XGBoost model exhibited a strong performance in predicting traffic operating speed. The model's predictions are close to the actual values, with minimal error, as evidenced by the low MAE and RMSE values for both the training and testing sets. The MAE for the training dataset as 0.108 and 0.137 for testing, indicating that the model's prediction deviated from the actual speed by an average of approximately 0.137 km/h. The R^2 value for the training set was 0.985, while the R^2 value for the testing set was 0.979 showing the model's prediction accuracy. This value indicates the percentage of efficiency of model 98.5% and 97.9% respectively for both training and testing, these values illustrates a strong relationship between the operating speed and the geometric properties of the highways (Yusuf *et al.*, 2016; Sun *et al.*, 2021).

For the training set, the COR values is 0.993 and testing set as 0.989. These values indicate the strength and direction of the linear relationship between the predicted and actual values. These high correlation values corroborate that the model's predictions are in good agreement with the observed data, indicating a strong positive relationship. The model's performance evidences its ability to accurately represent the intricate relationships between highway geometric properties and operating speed. The model's capacity to generalize

effectively to the testing data implies that it can be consistently employed to forecast operating speed in comparable scenarios. The results of this study have substantial implications for traffic management and highway.

Table 2: Performance Evaluation of Model

Metrics	R2	MAE	MSE	RMSE	COR
Training	0.985	0.108	0.913	0.116	0.993
Testing	0.979	0.137	0.021	0.144	0.989

IV. CONCLUSION

This research considered the influence of highway geometric in the prediction of traffic speed along a multilane highway in Nigeria through integration of Extreme Gradient Boosting (XGBoost) model for intelligent transportation system. The findings indicated that traffic operating speed prediction presents a robust and interpretable approach, considering road geometry, capturing complex relationships and incorporating probabilistic reasonings. It was discovered that the increase in shoulder width and pavement width results in an increase in operating speeds, which in turn enhances the safety and comfort of the driver. Furthermore, it was determined that roundabout intersections enhanced traffic flow, although traffic speed reduces at intersections. Road management agencies should consider the placement of traffic signs, speed

bumps, or pedestrian crossings to enhance road safety along routes with high interference.

XGBoost model provides a promising avenue for improving traffic speed prediction, ultimately contributing to safer and more efficient road networks. The findings of the study can serve as a basis for future highway design and policy-making, underscoring the significance of geometric features in the enhancement of the efficiency and safety of roadways. Although, XGBoost models provide positive results in predicting traffic operating speed, further research should be explored to evaluate the models using new datasets across other jurisdictions. In addition, evaluating and validating other ML models can be explored to predict traffic operating speed and improve road operation and effectiveness. Future research can focus on incorporating real-time data sources, such as GPS-based vehicle tracking and IoT sensors, to further enhance prediction accuracy and adaptability.

AUTHOR CONTRIBUTIONS

O. Bayode: Conceptualization, Methodology, Validation, Writing – original draft, Writing – review & editing. **O. Aiyeleokun:** Software, Data analysis. **A. A. Olanipekun & D.I. Idusuyi:** Formal analysis, Writing – review & editing. **E.O. Akande:** Writing – review & editing.

REFERENCES

- AASHTO. (2012).** American Association of State Highway and Transportation Officials. In A policy on geometric design of highways and streets. Washington, DC.
- AASHTO. (2014).** Highway Safety Manual. In A. A. (AASHTO). Washington DC.
- Abdelraouf, A., Abdel-Aty, M., & Yuan, J. (2022).** Utilizing attention-based multi-encoder-decoder neural networks for freeway traffic speed prediction. *IEEE Transactions on Intelligent Transportation Systems*, 23(8), 11960-11969.
- Ahmed, S., Hossain, A., Ray, S. K., Bhuiyan, M. I., & Sabuj, S. R. (2023).** A study on road accident prediction and contributing factors using explainable machine learning models: analysis and performance. *Transportation Research Interdisciplinary Perspectives*, 1-16.
- Amiri, A. M., Sadri, A., Nadimi, N., & Shams, M. (2020).** A comparison between artificial neural network and hybrid intelligent genetic algorithm in predicting the severity of fixed object crashes among elderly drivers. *Accident Analysis and Prevention* 138, 105468.
- Bayode, O., Aderinola, O.S., & Oluyemi-Ayibiowu, B. D. (2025).** Application of Machine Learning for Road Safety Modeling of Selected South-West Highway in Nigeria. *European Journal of Applied Science, Engineering and Technology*, 3(3), 202-213
- Bayode, O., Aderinola, O. S., Oluyemi-Ayibiowu, B. D., Okungade, O., & Opeyemi, D. (2024).** Impact of Geometric Properties of Two Lane Highway on Traffic operation and Safety. *UI Journal of Civil Engineering and Technology*, 6(1), 15-25.
- Berhanu, Y., Schroder, D., Wodajo, B. T., & Aliimayehu, E. (2024).** Machine learning for predictions of road traffic accidents and spatial network analysis for safe routing on accident and congestion-prone road networks. *Result in Engineering*, 23(102737), 1-22.
- Bouktif, S., Fiaz, A., Ouni, A., Alnaqbi, B., Alsereidi, F. S., & Alsereidi, F. A. (2020).** Bayesian optimized XGBoost model for traffic speed prediction incorporating weather effects. Fourth International Conference On Intelligent Computing in Data Sciences (ICDS), 1-7.
- Chen, M. M., & Chen, M. C. (2020).** Modeling road accident severity with comparisons of logistic regression, decision tree and random forest. *Information*, 11(5), 270.
- Chen, T., & Guestrin, C. (2016).** XGBoost: A Scalable Tree Boosting System. (pp. 1-8). San Francisco, CA, USA,: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.
- Dong, N., Huang, H., & Zheng, I. (2015).** Support vector machine in crash prediction at the level of traffic analysis zones: Assessing the spatial proximity effects. *Accident Analysis & Prevention* 82, 192-198.
- Ganan-Cardenas, E., Pemberthy-R, J. I., Suarez-Gomez, M. L., Ballesteros, J. R., & Branch-Bedoya, J. W. (2025).** Machine learning framework for predicting urban road speed profiles and uncertainty. *International Journal of Transportation Science and Technology*, 30-41.
- Imoh, U. U., & Rad, M. M. (2025).** Analysis and Prediction of Traffic Conditions Using Machine Learning Models on Ikorodu Road in Lagos State, Nigeria. *Infrastructures: MDPI*, 10, 1-24. Retrieved from <https://doi.org/10.3390/infrastructures10050122>
- Khanum, H., Garg, A., & Faheem, M. I. (2023).** Accident severity prediction modeling for road safety using random forest algorithm: an analysis of Indian highways [version 2; peer review: 2 approved, 1 approved with reservations]. *F1000Research*, 1-20.
- Khosbayar, A., Valluru, J., & Huang, B. (2021).** Multi-rate Gaussian Bayesian network soft sensor development with noisy input and missing data. *Journal of Process Control*, 105, 48-61.
- Lee, E. H. (2023).** Traffic Speed Prediction of Urban Road Network based on High Importance Links using XGBoost and Shapley Additive Explanation. *IEEE Access*, 1-10. doi:DOI 10.1109/ACCESS.2023.3324035
- Lee, E. H., Kho, S. Y., Kim, D. K., & Cho, S. H. (2021).** Travel time prediction using gated recurrent unit and spatio-temporal algorithm. *Proceedings of the Institution of Civil Engineers: Municipal Engineer*, 174(2), 88-96.
- Ma, D., Zhu, J., Song, X. B., & Wang, X. (2023).** Traffic flow and speed forecasting through a Bayesian deep multi-linear relationship network. *Experts system with Application*, 213, 119161.
- Ma, Z., Mei, G., & Cuomo, S. (2021).** An analytic framework using deep learning for prediction of traffic accident injury severity based on contributing factors. *Accident Analysis and Prevention*, 120, 106322.
- Min, J. H., Ham, S. W., Kim, D. K., & Lee, E. H. (2023).** Deep multimodal learning for traffic speed estimation combining dedicated short-range communication and vehicle detection system data. *Journal of the Transportation Research Board*, 2677(5), 247-259.

- Montella, A., Ona, R. D., Manuriello, F., & Riccardi, M. R. (2020).** A data mining approach to investigate patterns of powered two-wheeler crashes in Spain. *Accident Analysis and Prevention* 134, 105251, 1-15.
- Mystakidis, A., Koukaras, P., & Tjortjis, C. (2025).** Advances in Traffic Congestion Prediction; An Overview of Emerging Techniques and Methods. *Smart Cities (MDPI)*, 8(25), 1-37.
- Naallah, A. B., Ibitoye, A. B., & Subair, S. O. (2024).** Development of a Predictive Model for Road Traffic Accident Involvement in Southwest Nigeria: A Case Study of Ekiti, Osun, and Oyo State. *Journal of Applied Science and Environmental Management*, 3385-3390.
- Olugbemiro, A. (2019).** Road Traffic Accidents and Safety Challenges in Nigeria. *Journal of Transporton Safety.*, 337-342.
- Parsa, A. B., Movahedi, A., Taghipour, H., Derrible, S., & Mohammadian, A. K. (2020).** Toward safer highways, application of XGboost and SHAP for real-time accident detection and feature analysis. *Accidient Analysis Prevention*, 105405, 136.
- Roess, R. P., Prassas, E. S., & McShame, W. R. (2011).** *Traffic Engineering* (4th ed.). New York: Pearson Education.
- Semeida, A. M. (2013).** Impact of highway geometry and posted speed on operating speed at multi-lane highways in Egypt. *Journal of Advanced Research*, 4(1), 515-523. doi:<http://dx.doi.org/10.1016/j.jare.2012.08.014>
- Shiran, G., Imaninasab, R., & Khayamim, R. (2021).** Crash severity analysis of highways based on multinomial logistic regression model, decision tree techniques, and artificial neural network: A modeling comparison. *Sustainability*, 13(10), 5670.
- Sun, B., Sun, T., & Jiao, P. (2021).** Spatio-Temporal Segmented Traffic Flow Prediction with ANPRS Data Based on Improved XGBoost. *Journal of Advance Transportation*, 1-12.
- Tayebikhorami, S., & Sacchi, E. (2022).** Validation of Machine Learning Algorithms as Predictive Tool in the Road Safety Management Process: Case of Network Screening. *Journal of Transportation Engineering*, 1-30.
- Yang, Y., Wang, k., Yuan, Z., & Liu, D. (2022).** Predicting Freeway Traffic Crash Severity Using XGBoost-Bayesian Network Model with Consideration of Features Interaction. *Journal of Advanced Transportation*, 1-16.
- Yin, X., Fallah-Shorshani, M., McConnell, R., Fruin, S., Chiang, Y. Y., & Franklin, M. (2023).** Quantile extreme gradient boosting for uncertainty quantification. 20-32. Retrieved from <http://arxiv.org/abs/2304.11732>.
- Yusuf, I. T., Adeleke, O. O., Salami, A. W., & Ayanshola, A. M. (2016).** The Factors That Affect the Free Flow Speed on An Arterial In Ilorin, Nigeria. *Nigerian Journal of Technology*, 35(3), 473-480. Retrieved from <http://dx.doi.org/10.4314/njt.v35i3.2>.