AI-enabled Link Selection Framework for LTE Data Communication Towards Collision-Managed Traffic

Bhimani Yatra Amitbhai*, Bhayani Sahi†, Keertan Amit Parikh‡, Rajesh Gupta§, Sudeep Tanwar¶
*†‡§¶Department of Computer Science and Engineering, Institute of Technology, Nirma University, Ahmedabad, India
Emails: *22bce039@nirmauni.ac.in, †22bce037@nirmauni.ac.in, ‡21bce117@nirmauni.ac.in
§rajesh.gupta@nirmauni.ac.in, ¶sudeep.tanwar@nirmauni.ac.in,

***Abstract*—Developing technologies, such as 5G networks for wireless communication, often leverage the concept of V2X communication. For any communication, the path selected for it plays a vital role as it needs to transmit data between interconnected devices in less time and efficiently. This research discusses selecting an efficient communication channel in cellular infrastructure, i.e., C-V2X. In this exploration, two base stations, namely Vodafone and Deutsche Telekom, taken from the Berlin dataset help to deliver an approach for selecting a link based on their performance. The research is done by training diverse ML models such as Random Forest, XGBoost, and Elastic net Model on the dataset. These are then evaluated using their respective RMSE and R-square values that help in better interpretation of the base station link. It aims to select a convenient path to transmit data at low latency, provide reliable and secure communication, and exhibit dynamic link adjustment in changing environmental circumstances.**

***Index Terms*—V2X communication, link selection, cellular, machine learning**

I. Introduction

Link selection in V2X refers to selecting an appropriate channel to transmit data between two inter-connected nodes. The main objective of link selection is to provide reliable data communication by considering factors such as latency, data rate, throughput, bandwidth, and other environmental factors. This study aims to choose the most reliable link for communication between multiple base stations based on their reach and efficiency in data transmission. According to [1], around 1.19 million people lead to death from road accidents globally. V2X communication can help to reduce road accidents by giving expeditious feedback on the changing road conditions. It also facilitates traffic management by providing real-time traffic updates which helps vehicles select less congested routes.

The traditional solutions, such as improvements in the infrastructure, enforcement of laws, and safety regulations in vehicles taken to solve this problem, are reactive rather than proactive. These solutions do not address the problem as they do not leverage real-time data transfer of changing road conditions and driving behaviors. They are also time-consuming and costly. In [2], the Marouan *et al.*, relied on beam selection, which encompasses the most optimal and efficient beamforming direction for data transmission between communication entities. Q-REDD's method of stable matching-based routing, D2D communication, and emergency message dissemination techniques offers scalability and adaptability in dynamic V2X environments Saleh *et al.* [3]. Using a multi-agent actor-critic approach, it can adapt to changing V2X data traffic patterns and adjust resource allocation parameters based on real-time conditions, according to Anupama *et al.* [4].

The paper [5] by Bharatwaja *et al.* included a broader spectrum of AI technologies and communication standards in autonomous driving. The optimization scheme and resource management principles given in Ref. [6] for 5G networks support V2X communication indirectly and express link selection for V2X communication by reducing noise, enhancing network performance, and minimizing resource utilization. The overall performance can be enhanced by minimizing the number of active RRHs based on traffic demand and uplink bandwidth constraints, leading to more effective link selection. In their paper, Sohan *et al.* [7] addressed the problems of LTE-based vehicle-to-everything communication. It mainly talked about resource allocation, which is important for selecting links for users, security, and the need for a physical layer structure for LTE V2X communication by supporting link selection. The 5G for V2X communication Ref. [8] highlighted key enabling technologies for next-gen V2X communication. It focused on the need for an extraordinary V2X communication network with the capacity to facilitate hyperfast, ultra-reliable, and low-latency information exchange. The real-time data exchange between the vehicles through wireless networks with speed, direction, and position connects to link selection for V2X communication by Mohammed *et al.* [9]. Their study provides insights into 5G communication technologies and the technological aspects of dedicated short-range communications (DSRC) and cellular vehicle-to-everything (C-V2X) communication.

In this study, it has incorporated dimensionality reduction technique, and principal component analysis, which reduces the dataset's dimensions, providing valuable insights and enhanced visualization. It helps in the dimensional reduction of the features by grouping features with similarities in one principal component. Many papers lack this technique, but this research helped in feature extraction and noise reduction, which led to a focus on the significant features. Further, the missing values of the independent features of the data set [10] were handled concerning the target variable latency. The latency range was found, and then equal groups were made based on their percentile distribution. Considering one group of latency, data on all the independent features lying in that range was taken, and then mean imputation was performed on it. The missing values in that stratum were replaced by the mean. This strategy upholds context-relevant services and safeguards the dataset's integrity. Regularization of data is also done to prevent the model from overfitting. If the model overfits, it only learns the data fluctuations from the dataset used for training and does not provide an efficient result on unseen data. Performing this technique makes the research more generalized and also enhances its reliability.

A. Research Contributions

The research contributions are as follows:

- Extensive data processing has been performed on data from diverse driving environments with multiple operators, including handling missing values and outliers, principal component analysis for feature reduction, and normalization.
- We propose a framework for efficient communication channel selection (link selection) using ML to perform low-latency data transmission. It requires training multiple ML models, such as random forest, XGBoost, and linear regression, along with elastic-net regularization, to predict and compare data transmission latency in the different network operators.
- The model's robustness is ensured using kfold cross validation and its performance is assessed using evaluation metrics including rmse and r-square values.

B. Organization of the Paper

The rest of the paper is organized as follows: Section II discusses the system model and problem formulation, and Section III presents a detailed description of the proposed system approach. Subsequently, Section IV offers an assessment of the proposed approach's performance. Finally, section V shows the research conclusion and prospects.

II. System Model and Problem Formulation

The proposed system focuses on link selection, a crucial aspect of V2X communication. For this purpose, the data is collected by a vehicle movement measurement campaign in Berlin by framing a route that includes urban areas, parks, tunnels, and highways. This campaign was carried out with two driving modes: platoon and 2x2. Measurement tools are placed in each vehicle, which helps test, analyze, and measure the data related to the cellular network. A wireless communication technology is transmitted to the servers following the data-gathering process. The TCP/IP protocol is used at the transport layer to transmit information to the server. The data is pre-processed at the server, where the missing values are handled, and normalization is done if needed. After pre-processing comes data processing, where the patterns in the data are analyzed, and the ML algorithms are involved. The pre-processed and processed data are stored in the server's database. The server sends Data to base stations when the base station requests it. The problem that arises in the transfer is data security. To solve this, the data is encrypted, and to optimize the transfer process, It is compressed before transmitting. Special protocols are being used to transmit the data from the server to the base station for reliable and efficient data transfer. In the current generation, 5G, the latest technology in the industry, is employed to transmit data at increased data rates and reduced latency. The data received at the base station is decoded and processed, then transferred to the connected vehicle in the coverage range of the base station.

During the measurement, the vehicles will be continuously moving, so collecting data at every moment is difficult. The solution is sampling, which means collecting data at a specific interval. The sampling interval helps collect the data efficiently and also helps maintain the uniformity of the data. The given formula determines the sampling interval and is denoted by α in Eq. 1.

$$\alpha = \frac{\mathrm{D}}{\mathrm{M}} \tag{1}$$

where,
(D) denotes the total measurement duration and (M) is the total number of measurements.

The distance traveled by vehicles during the campaign is symbolized as β in Eq. 2.

$$\beta = N \times RL \tag{2}$$

where,
(N) denotes the total number of rounds taken. (RL) is the route per round length.

For identifying the quality of a link, it is important to calculate the signal-to-noise ratio (SNR). A higher SNR value indicates less noise interference and an enhanced quality link. For inspecting the received signal strength aligned with noise, SNR is calculated as Eq.3.

$$\mathrm{SNR} = \frac{\text{Signal power}}{\text{Noise power}} \tag{3}$$

This research is for the application in real-time; for this objective, it is significant to calculate the data rate value. It helps determine how rapidly the data can be transmitted. For this approach, it is denoted by θ in Eq. 4.

$$\theta = \frac{\mathrm{S}}{\mathrm{T}} \tag{4}$$

where,
(S) denotes the data received. (T) denotes duration for data transmission.

The packet error ratio (PER) is evaluated to find the error rate. Lower values of PER mean fewer packets with errors, and the link has a high data throughput. It is denoted by λ in Eq. 5.

$$\lambda = \frac{\text{Number of Corrupted Packets}}{\text{Total Number of Packets}} \tag{5}$$

The link selected can transmit data between two vehicles effectively without any loss while maintaining the signal strength and quality over a limited range. γ symbolizes the link range and is calculated as Eq. 6

$$\gamma = \mathrm{F}(\beta, \mathrm{SNR}, \lambda) \tag{6}$$

where (F) represents the function that calculates the link range based on the three parameters. β is the distance calculated in Eq. 2, SNR ratio calculated in Eq. 3, In Eq. 5, λ represents the packet error ratio.

By training the machine learning (ML) algorithms on the traits of the network layer, the latency can be predicted. It is denoted by δ in Eq. 7.

$$\delta = M_{\mathrm{NLC}} \tag{7}$$

where (M) is the ML algorithm trained on network layer characteristics (NLC), such as IP addressing, error identification and resolution, network congestion, routing protocols, and more. To determine the data transmission rate from the source to the destination, Throughput needs to be assessed. It helps to inspect and analyze the performance of the network. It is denoted by η in Eq. 8.

$$\eta = T_{\mathrm{ALC}} \tag{8}$$

where (T) is the ML algorithm trained. ALC represents the application layer characteristics such as data authentication and authorization, data compression, resource sharing, and application protocols such as HTTP and FTP.

III. Proposed Approach

This section presents the layer framework for AI-based link selection for V2X communication. It has four layers: data acquirement, data pre-processing, training, and communication marginal. The following subsections describe each layer.

A. Data Acquirement Layer

The Berlin dataset has been collected from drive tests around West Berlin, along with some routes. The route included highways, parks, residential areas, and tunnels. The route length is 17.2 km, and in one run, it takes 45 minutes for a round on a weekday morning. In 3 days, the Berlin dataset developers have driven 17 clockwise rounds along with the measurements of the route, with 4 cars operating on two of these driving modes. One is Platoon, and the other is 2x2. Each car is given a unique UL/DL profile for each round used for data transmission. Also, unique port numbers are assigned to each car, which helps in data pre-processing. Sensors like an accelerometer are installed to check the car's acceleration, a gyroscope is employed for assistance in navigation and orientation, and GPS is installed for location. Data is exchanged between the car and the server using LTE by employing the corresponding measurement equipment in each car. Among the four cars, two are connected to Vodafone, and the other two are connected to Deutsche Telekom. In the rest of the paper, it will be denoted as Operator 1 and Operator 2, respectively.

With the DME installed in each car, there is an antenna, modem, and GPS receiver, which help in communication with the cellular network. It also uses applications such as MobileInsight, which helps with cellular network performance; TCPdump, which is used to get network traffic data; and Iperf, which is used for maximum bandwidth performance. Firstly, DME transfers the data collected by the sensors to the base stations in the coverage area over LTE. Once the data is received at the base station, it is further transferred to the destination server through the core network for further data processing.

B. Data pre-processing layer

The Berlin dataset [10] contains many missing values that must be handled before making any prediction. Firstly, five features with the most missing values were taken and handled. Take 5 ranges of the target variable (latency), separate the rows for each range, take a mean of those rows of that particular feature, and replace them with the missing values present for the same latency range. The same methodology will be applied to both datasets. Principle component analysis (PCA) applied to the dataset was found to be important and affected features for link selection. It helps in reducing features for a better prediction process. For the rest of the column's missing values, handling typically takes a mean of all the present values and replaces the missing ones with the mean. The dataset is normalized because it has to be scaled down between a common range to ensure that all the features present in each dataset contribute equally to the predictions. After the pre-processing, the data is further analyzed using ML algorithms.

$$\mathrm{cov}(X, Y) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \tag{9}$$

where,

n is the number of observations.

x_i and y_i are individual observations of variables X and Y.

$\bar{x}$ and $\bar{y}$ are the sample means of variables X and Y, respectively.

The covariance matrix is used in PCA because it gives data about the variance and relationship between the features in the dataset.

$$\sum v = \lambda v \tag{10}$$

where,

v is an eigenvector.

λ is an eigenvalue.

Eigenvalues show the quantity of variance explained by each PCA. To prevent overfitting, CV is being used (cross-validation) with 5 splits, dividing the whole training dataset into 5 parts, followed by different training and validation data partitions with divergent R squares, which will then be averaged to calculate the final R square value.

C. LatencyModelTrain Layer

Three models, namely, random forest, XGBoost, and linear regression, were trained on the dataset. The random forest consists of tree-like structures in which every node shows a test on an attribute, every branch shows the answer to the test, and

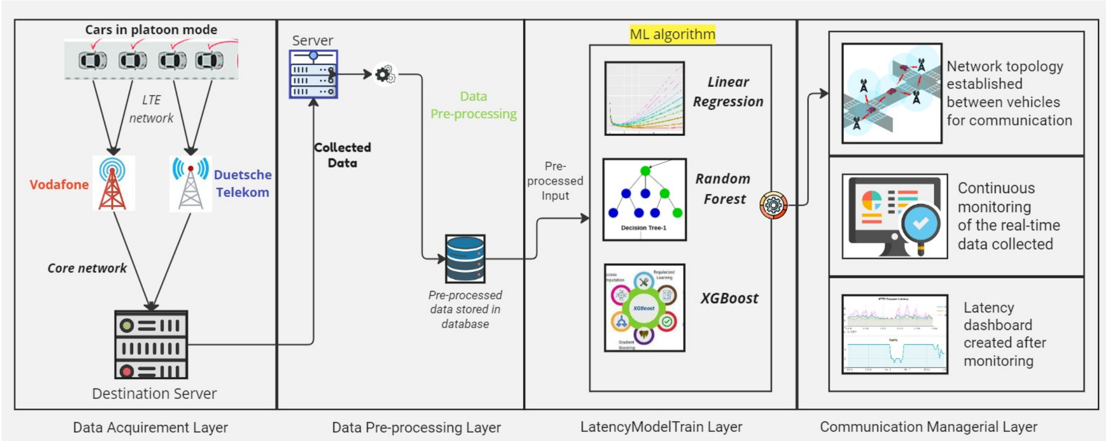


Fig. 1: System model

every leaf node shows a numeric value in regression. Secondly, the XGBoost is an advanced part of gradient boosting, in which many weak learners are combined to generate a strong learner. In this way, the errors of each model can be corrected sequentially. It works on a specific function that needs to be reduced during training. In regression, the main goal is to reduce MSE. Linear regression is a method that consists of one target variable with multiple features. It is the most commonly used technique in machine learning. In the case of multiple-featured linear regression, it can be represented by a hyperplane.

Regularization prevents over-fitting, making the coefficient of less important features zero and facilitating feature selection. Regularization also helps the ML model not to be highly influenced by the outliers, and specifically, elastic-net regularization helps in flexibility and adaptability as it combines both Lasso (L1) and Ridge (L2) regularization. Eventually, implemented all these techniques to prevent overfitting and trained all the models again; additionally identified the R-square and RMSE values of the models.

$$\text{ElasticNet}(\beta) = \text{L} + \lambda_1 \left(\alpha \sum_{j=1}^{p} |\beta_j| \right) + \lambda_2 \left(\frac{1}{2}(1-\alpha) \sum_{j=1}^{p} \beta_j^2 \right) \tag{11}$$

where,
L = loss function (shows the original loss function being improved (e.g., MSE for linear regression))
λ_1 and λ_1 are the regularization parameters.
β are the coefficients of the features. α is the mixing parameter between Lasso and Ridge regularization.

D. Communication managerial layer

Once the regression models train the data, the communication managerial layer will be introduced, which describes how the data will be further transmitted to other entities. A network topology connecting vehicles, base stations, infrastructure, servers, and pedestrians is established for the communication process. The connection is established between two entities considering various factors such as distance, signal strength, throughput, etc., which can be either a point-to-point or a hop-to-hop connection. The network topology helps route the data from one source vehicle to the destination point.

The most significant segment of the task involves sorting out the path with the least congestion and which has the shortest distance during routing. This model continuously analyzes and elucidates the real-time data, thus efficiently monitoring and achieving the aforementioned task. Once the path is selected, the packet is sent through it and continuously monitored to see if it follows it. The security of data is a crucial aspect of the transmitting process. If any cyber attacks on data occur during the process, the monitoring device can detect them and also send alerts to them.

In V2X communication, secured transmission is the supreme criterion; to achieve it, the data should be transferred at low latency and the communication should be reliable. For this purpose, a latency monitoring dashboard is maintained, ensuring the data gets transferred at a very low latency. It

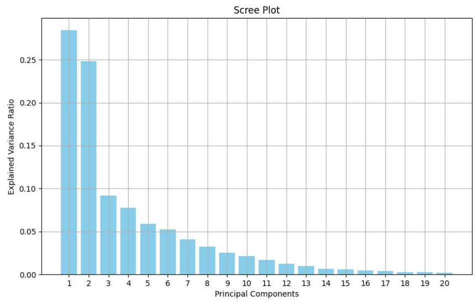
Algorithm 1 Working of machine learning XGBoost algorithm for link selection.

1: Data collected from Berlin: df $= \{(x_i, y_i)\}_{i=1}^{N}$, where $x_i \in \mathbb{R}^d$ and $y_i \in \{1, \ldots, C\}$
2: Prediction: y_{pred}
3: Apply PCA to extract the most relevant features from the dataset.
4: Split dataset into training and testing sets: X_train, X_test, y_train, y_test
5: Identifying and removing the outliers using
- $\mathcal{Q}_1 =$ df.quantile(0.25)
- $\mathcal{Q}_3 =$ df.quantile(0.75)
- IQR $= \mathcal{Q}_3 - \mathcal{Q}_1$
- Outliers $=$ (df $< (\mathcal{Q}_1 - 1.5 \times$ IQR$)) \mid$ (df $> (\mathcal{Q}_3 + 1.5 \times$ IQR$))$

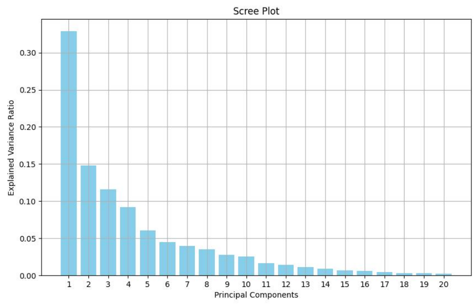
6: Extract the independent and dependent variables.
7: Feature scaling using StandardScaler.
8: Convert data to Dmatrix format.
9: Define parameters for XGBoost like: objective, eval_metric, eta, max_depth, subsample, colsample_bytree, reg_alpha, reg_lambda.
10: Define the number of boosting rounds.
11: Perform KFold Cross Validation using equation 9 and get predictions.
12: Calculate RMSE and R-square using eq 12 and 14 between the training and testing data respectively.
13: Final RMSE and R-square value shows the efficiency of the model.

can also help make proactive decisions by analyzing the historical latency data and finding solutions. By continuously analyzing the dashboard data, it incorporates changes such as routing algorithms to be used, the addition of a new network component, updates in the traffic, etc.

IV. Performance Evaluation



(a) Operator 1



(b) Operator 2

Fig. 2: PCA analysis

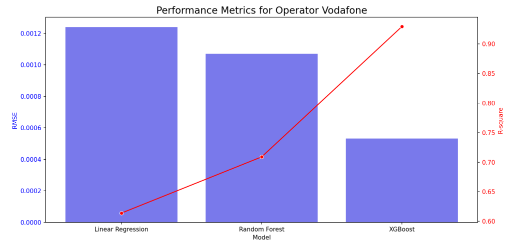
In Fig. 2, scree plots are plotted about Principal Component Analysis (PCA). It determines the main features that affect the target variable (latency) the most. A scatter plot gives data about insignificant features and helps to reduce them.

Fig. 3 visualizes the R-square and RMSE values for the various models utilized. The graphs show that the correlation coefficient, r-square value, is consistently rising, approaching 1. As explained above, the closer to 1 value of R-square is better. Also, graphs indicate the RMSE values for each model, and as it continuously decreases, it shows that the error is lower and the model is better.

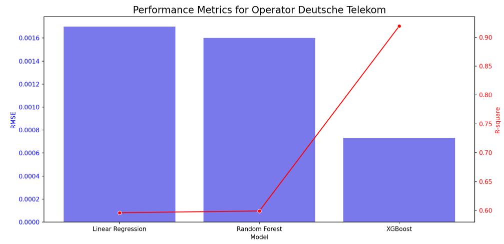
TABLE I: Performance Metrics of Models

Model	**RMSE**	**R-square**
Linear Regression	0.00124	0.614
Random Forest	0.00107	0.709
XGBoost	**0.000533**	**0.9292**

Table I is the performance matrix [evaluation matrix]. The matrix indicates the numeric values of each model trained on the Berlin dataset for link selection for V2X communication. The R-square value for operator 1 with the XGBoost regressor is 0.9292, and the RMSE is 0.000533. It emphasizes that the XGBoost regressor fits the data set well.



(a) Operator 1



(b) Operator 2

Fig. 3: RMSE and R-square values of various ML models

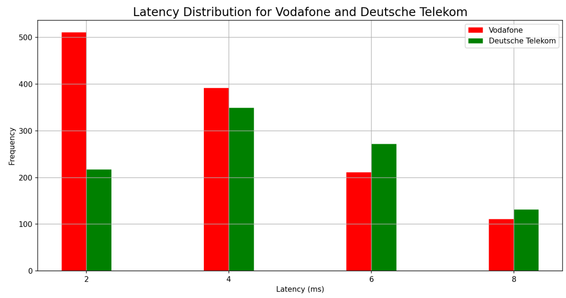


Fig. 4: Latency distribution

TABLE II: Performance Metrics of Models

Model	**RMSE**	**R-square**
Linear Regression	0.0017	0.596
Random Forest	0.0016	0.599
XGBoost	**0.000732**	**0.9193**

Table II shows all the values of operator 2. The values for the same are 0.9193 and 0.000732, respectively, for the XGBoost model.

$$\text{MSE} = \frac{\sum_{i=1}^{N}(P_i - P_i')^2}{N} \tag{12}$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(P_i - P_i')^2}{N}} \tag{13}$$

Once the model is successfully trained, it will make further predictions on the testing dataset to evaluate the model's performance. This is achieved by comparing the testing dataset's actual values with the model's predicted values. Then will use the evaluation method of RMSE, which provides an indicator for the difference between the predicted and actual values of the target variable. RMSE is an error, so if there is any non-zero difference, that would be calculated, and a mean would be taken of these differences. The square root of that mean would be the final value of the RMSE. Eq. 13 is about the RMSE [11].

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(P_i - P_i')^2}{\sum_{i=1}^{N}(P_i - \bar{P})^2} \tag{14}$$

where,
P_i' is the mean value of the predicted values.
P_i is the predicted value of the target variable for the ith data point.
$\bar{P}_i$ is the mean value of the target variable.

The R-square value is beneficial in comparing each model by providing a standardized measurement of how well the data fits. Eq. 14 is the R square formula, which balances model simplicity and explanative potency [12].

Choose the prediction's R-square and RMSE values rather than the MSE, MAE, or MAPE. RMSE gives an error metric in equal units as the target variable, making it more preferable and interpretable. MSE is more applicable for model training optimization as it gives gradients but may not ease error interpretation. RMSE is better than MAE because MAE doesn't give any insights into error distribution and is also not sensitive to deviations like RMSE. For R-square and adjusted R-square, R-square directly measures how well the features other than the target variable explain the variability of the target variable. On the other hand, adjusted R-square takes the number of predictors in the model and removes the inclusion of non-relevant variables that do not enhance the model's performance. Finally, for these reasons, RMSE and R-square values are easier to find and use to check the model's effectiveness.

The reason behind opting for the XGBoost is that it fits the dataset perfectly with the values of 0.9292 and 0.000533 for Vodafone infrastructure, and it doesn't show overfitting. Linear regression and random forest fit the dataset well. However, a closer-to-unity R-squared value demonstrates that the independent variables explain a larger proportion of the variance in the target variable. As a result, a higher R-squared value shows better predictive performance, which indicates that the model records more of the underlying patterns in the data. In regression, R-squared measures the variance in the target variable that is predictable from the independent variables, and for those reasons, the XGBoost is chosen.

Fig 4. shows the operators' latency distribution graph. This bar graph shows that operator 1 has a higher frequency at lower latency values, whereas operator 2 has a higher frequency at higher latency values. This concludes that operator 1 has better performance.

V. CONCLUSION

This paper concludes that link selection is important for applications such as collision avoidance, which makes traffic flow smoother. It helps in traffic management, prevents roadside accidents on time or before time, and aids in a quick response process. For the link selection process for V2X communication, latency as a criterion is selected, and it is trained concerning the other features affecting the communication. For this approach, the data has been trained using different models like linear regression, random forest, and XGBoost. Among all these models, XGBoost performs the best by giving values of R-square 0.9292 and RMSE 0.000533. On further examination of the values of latency, it can be observed that a lower value of latency shows better performance; therefore, it leads to faster data transmission and diminishes the communication delay, and additional work can be done to facilitate more features under consideration to optimize the model further.

The future scope of this research lies in extending the potential of the study to diverse environments and adapting to dynamic situations such as the unavailability of complete data, traffic patterns of urban and rural areas, and infrastructure designs. Applying furtherance increases the reliability, safety, and productivity of V2X communication, which eventually contributes to modern Internet communication.

REFERENCES

[1] [Online]. Available: https://www.researchgate.net/publication/372824335

[2] M. Mizmizi, F. Linsalata, M. Brambilla, F. Morandi, K. Dong, M. Magarini, M. Nicoli, M. N. Khormuji, P. Wang, R. A. Pitaval, and U. Spagnolini, "Fastening the initial access in 5g nr sidelink for 6g v2x networks," *Vehicular Communications*, vol. 33, 1 2022.

[3] S. A. Alghamdi, "Cellular v2x with d2d communications for emergency message dissemination and qos assured routing in 5g environment," *IEEE Access*, vol. 9, pp. 56 049–56 065, 2021.

[4] A. Hegde, R. Song, and A. Festag, "Radio resource allocation in 5g-nr v2x: A multi-agent actor-critic based approach," *IEEE Access*, vol. 11, pp. 87 225–87 244, 2023.

[5] B. Namatherdhala, N. Mazher, G. K. Sriram, and G. K. Sriram, "Uses of artificial intelligence in autonomous driving and v2x communication," *www.irjmets.com @International Research Journal of Modernization in Engineering*, 2022. [Online]. Available: www.irjmets.com

[6] R. H. Hwang, F. Marzuk, M. Sikora, P. Cholda, and Y. D. Lin, "Resource management in ladns supporting 5g v2x communications," vol. 2020-November. Institute of Electrical and Electronics Engineers Inc., 11 2020.

[7] S. Gyawali, S. Xu, Y. Qian, and R. Q. Hu, "Challenges and solutions for cellular based v2x communications," *IEEE Communications Surveys and Tutorials*, vol. 23, pp. 222–255, 1 2021.

[8] M. Noor-A-Rahim, Z. Liu, H. Lee, M. O. Khyam, J. He, D. Pesch, K. Moessner, W. Saad, and H. V. Poor, "6g for vehicle-to-everything (v2x) communications: Enabling technologies, challenges, and opportunities," *Proceedings of the IEEE*, vol. 110, pp. 712–734, 6 2022.

[9] M. Alsudani, T. Ozturk, and T. Öztürk, "Wireless communication between vehicles: Exploring the potential of v2v and v2x communication for improved efficiency, safety, and sustainability." [Online]. Available: https://www.researchgate.net/publication/372824335

[10] R. Hernangómez, P. Geuer, A. Palaios, D. Schäufele, C. Watermann, K. Taleb-Bouhemadi, M. Parvini, A. Krause, S. Partani, C. Vielhaus, M. Kasparick, D. F. Külzer, F. Burmeister, F. H. P. Fitzek, H. D. Schotten, G. Fettweis, and S. Stańczak, "Berlin v2x: A machine learning dataset from multiple vehicles and radio access technologies," 12 2022. [Online]. Available: http://arxiv.org/abs/2212.10343http://dx.doi.org/10.1109/VTC2023-Spring57618.2023.10200750

[11] T. O. Hodson, "Root mean square error (rmse) or mean absolute error (mae): when to use them or not," 2022. [Online]. Available: https://doi.org/10.5194/gmd-2022-64

[12] P. K. Ozili, "The acceptable r-square in empirical modelling for social science research," *SSRN Electronic Journal*, 6 2022.