Ain Shams Engineering Journal 17 (2026) 104102

Contents lists available at ScienceDirect

Ain Shams Engineering Journal

journal homepage: www.sciencedirect.com

Full length article

Environment-aware V2V communication: A machine learning framework for real-world propagation classification

Nuğman Sağır [a, b,*], Zeynep Hasırcı Tuğcu [c]

[a] Karadeniz Technical University, Graduate School of Natural and Applied Science, Trabzon, 61080, Türkiye
[b] Karadeniz Technical University, Faculty of Engineering, Department of Electronics and Communications Engineering, Trabzon, 61830, Türkiye
[c] Karadeniz Technical University, Department of Electrical and Electronics Engineering, Trabzon, 61080, Türkiye

ARTICLE INFO

Keywords:
Vehicle-to-vehicle communication
Machine learning
Environment classification

ABSTRACT

This study investigated the use of machine learning techniques for the classification of the propagation environment in vehicle-to-vehicle (V2V) communication systems under varying environmental conditions, to improve communication reliability and energy efficiency. To do this, field measurements were first conducted in Trabzon and Gümüşhane, Turkey, for different environments, including rural, highway, suburban, and urban environments, to analyze the propagation characteristics. Then, the classification performance of five machine learning models—K-Nearest Neighbors (KNN), Random Forest (RF), Artificial Neural Networks (ANN), Support Vector Machines (SVM), and AdaBoost—was evaluated and compared. Key input features, including distance, path loss, obstacle types, modulation schemes, and weather conditions, were weighted and utilized for classification. Here, the KNN model achieved superior performance with an accuracy of 99.84%, precision of 99.83%, recall of 99.84%, and an Area Under the Receiver Operating Characteristic Curve (AUC-ROC) of 99.97%. RF and ANN models also demonstrated high AUC-ROC values, underscoring their robustness in complex propagation environments. To understand the contribution of individual features to the classification performance, an Out-of-Bag Permutation Feature Importance analysis was performed. The results indicated that weighted building obstacles were the most influential feature, accounting for 20.09% of the total feature importance. These findings highlight the potential of adaptive power adjustments based on real-time environmental classification to improve both the reliability and efficiency of V2V communication systems, thereby supporting the development of safer and more efficient intelligent transportation networks.

1. Introduction

Vehicle-to-vehicle (V2V) communication has emerged as a cornerstone of intelligent transportation systems to enhance road safety, traffic efficiency, and vehicular automation. Recent advancements in path loss modeling for V2V communications have provided crucial insights into channel characteristics across diverse environments, such as highways, urban, and suburban areas [1]. These studies are complemented by research on Dedicated Short Range Communications (DSRC) and IEEE 802.11p, which have shown reliability challenges under non-line-of-sight conditions [2]. Meanwhile, significant developments in V2V communication protocols for cooperative collision-warning systems have demonstrated reduced latency and improved emergency message dissemination [3]. Furthermore, comprehensive reviews have highlighted the integration of V2V technologies with autonomous driving systems, emphasizing applications such as adaptive cruise control and lane-change assistance [4]. However, challenges, such as high mobility, dynamic channel conditions, and data security, underscore the need for innovative solutions [5] (Table 1).

1.1. Literature review

In recent years, machine learning techniques have been widely employed in classification tasks related to vehicular communication because of their ability to learn complex patterns, adapt to diverse environments, and enhance predictive accuracy. Various studies have explored different aspects of classification within vehicular networks. For instance, deep learning models, such as Convolutional Neural Networks (CNNs), have been utilized for vehicle classification under challenging conditions, achieving high accuracy [6]. Similarly, Long Short-Term Memory (LSTM) and Transformer models have been employed to classify vehicle position classes in V2V communications, demonstrating

* Corresponding author at: Karadeniz Technical University, Graduate School of Natural and Applied Science, Trabzon, 61080, Türkiye.
Email addresses: nugmansagir@ktu.edu.tr (N. Sağır), zhasirci@ktu.edu.tr (Z. Hasırcı Tuğcu).

https://doi.org/10.1016/j.asej.2026.104102
Received 14 April 2025; Received in revised form 12 July 2025; Accepted 5 March 2026

2090-4479/© 2026 The Author(s). Published by Elsevier B.V. on behalf of Faculty of Engineering, Ain Shams University. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Table 1
Literature review of classification-based studies.

Reference	Methods	Environment	Study area	Output parameter	Performance metric
[6]	CNN (ResNet)	Various	V2X Communication	Vehicle Type	$Accuracy = 99.68\%$
[7]	LSTM, Transformer	Highway	V2V Communication	Vehicle Position Classes	$Accuracy = 93\%$
[8]	SVM, RF	Urban, Suburban	Driving Environment Classification	Driving Environment	$Accuracy = 93\%$
[9]	Neural Networks, Matrix Classification	Highway, Urban	V2V, I2V, and V2I Communication	Rerouting Decisions	Improved rerouting efficiency by 15%
[10]	K-Nearest Mean	Rural	Land Use Classification	Land Use Classification	$Accuracy = 88\%$
[11]	CNN, Deep Learning	Satellite Images	Path Loss Prediction	Channel Parameters	$Accuracy = 88\%$
[12]	CNN, Hybrid Regression	Various	Path Loss Prediction with Images	Path Loss Prediction	$Accuracy = 88\%$
[13]	XGBoost	Various (SRM2KTR Dataset)	Traffic Monitoring	Vehicle Type	$Accuracy = 98.81\%$
[14]	BiGAN (Bidirectional GAN)	Vehicular Ad-Hoc Network	Security in V2X Communication	Anomaly Detection	$Accuracy = 92.15\%, F1 = 96.10\%$
[15]	Active Learning (AL)	Connected Vehicles (V2V)	Event Classification in V2V Communication	Event Classification	Improved classification accuracy with reduced bandwidth usage
[16]	Decentralized Federated Learning	Urban Roads (V2V)	Road User Classification in V2X Networks	Road User Classification	Improved accuracy with minimal bandwidth
[17]	DSRC-based Target Classification	Urban Roads (V2P)	Pedestrian Safety Systems	Pedestrian Safety	Enhanced collision detection with low latency
[18]	Memory and Learning-based Semantic Communication	Wireless Channels	Semantic Communication in 6G	Text/Image Classification	$Accuracy = 92\%$ (text), $Accuracy = 97.1\%$ (images)
[19]	Joint Detection and Modulation Classification (JDM)	Wireless Communication	Signal Processing for Modulation Detection	Modulation Type	$Accuracy = 95\%$ (low SNR)
[20]	Deep Learning-based Driver Classification	Vehicular Ad-Hoc Network (VANET)	Driver Trust Scoring in V2V Networks	Driver Trust Score	$Accuracy = 96.7\%$

significant performance improvements in highway environments [7]. Additionally, machine learning techniques, such as Support Vector Machines (SVM) and Random Forest (RF), have been used for driving environment classification, supporting eco-driving strategies by accurately identifying different driving conditions [8]. Some studies have focused on the application of neural networks and matrix classification techniques in V2V, Infrastructure-to-Vehicle (I2V), and Vehicle-to-Infrastructure (V2I) communications to improve traffic management, with models demonstrating notable enhancements in rerouting efficiency [9]. Moreover, land use classification models have been applied in rural areas, achieving reliable classification accuracy using K-Nearest Mean methods [10]. Beyond the classification of driving environments and vehicle-related attributes, machine learning approaches have been integrated into communication system modeling, particularly for path loss prediction. Studies have generally utilized deep learning models trained on satellite images to estimate channel parameters such as the path loss exponent and shadowing factor, achieving substantial accuracy in diverse environments [11]. Convolutional Neural Network (CNN)-based hybrid regression models have also been proposed for predicting path loss under different environmental conditions and have shown promising results [12]. In the context of traffic monitoring, large-scale datasets have been used to classify vehicles under varying environmental conditions, and ensemble learning models such as eXtreme Gradient Boosting (XGBoost) have achieved high accuracy [13]. On the other hand, security in vehicular networks has also been a significant research focus, with models based on bidirectional Generative Adversarial Networks (BiGAN) effectively detecting anomalies in Vehicle-to-Everything (V2X) communication with notable classification performance [14]. In addition to these advancements, active learning frameworks have been developed to enhance the classification accuracy of connected vehicle networks while optimizing bandwidth usage [15]. Decentralized federated learning approaches have further contributed to the classification of road users in V2X networks by selectively sharing model parameters, thereby reducing communication overhead while maintaining performance [16]. Similarly, Dedicated Short-Range Communication (DSRC)-based classification techniques improve pedestrian safety in urban Vehicle-to-Pedestrian (V2P) environments by enabling rapid target classification and collision detection [17]. Beyond vehicular communication, memory- and learning-based semantic communication frameworks have been proposed for efficient text and image classification in Sixth-Generation (6G) wireless networks, demonstrating high accuracy while optimizing system performance [18]. Furthermore, joint detection and modulation classification techniques have been introduced to improve modulation recognition accuracy in low-Signal-to-Noise Ratio (SNR) environments [19]. Deep learning-based driver classification frameworks have also been developed within Vehicular Ad-hoc Networks (VANETs) to assess driver behavior and compute trust scores, ultimately enhancing road safety by identifying unreliable drivers and improving communication efficiency [20].

1.2. Motivation and contribution

Propagationenvironmental classification plays a crucial role in optimizing V2V communication systems by enabling adaptive strategies to ensure communication reliability across different scenarios. V2V communication occurs in highly dynamic and heterogeneous environments, where factors such as obstacle density, vehicle mobility, terrain type, and weather conditions significantly influence signal propagation and communication quality. For example, urban environments with dense buildings and high vehicular traffic exhibit vastly different propagation characteristics compared with rural or highway settings, where line-of-sight (LOS) communication is more dominant. These variations necessitate adaptive mechanisms that dynamically adjust the transmission parameters, such as power control and modulation schemes, to maintain efficient and reliable communication. Although GPS-equipped vehicles can determine their geographic location, GPS alone cannot capture real-time propagation characteristics such as temporary obstacles, dynamic traffic flow, or weather-induced signal degradation. Therefore, machine learning-based classification using signal and environmental features offers a more adaptive and context-aware approach to improve V2V communication reliability. Although numerous studies have explored different aspects of V2V communication, including vehicle classification [6], modulation detection [19], and event classification [15], relatively few have focused on environmental classification. Some existing studies have examined specific environments, such as rural

areas [10] or urban and suburban settings [8]; however, they typically address single-environment classification rather than a comprehensive multi-environment approach. Comprehensive studies covering multiple driving environments—rural, highway, suburban, and urban—are notably limited in the existing literature. This gap is critical because precise environmental classification can allow V2V systems to optimize communication parameters in real-time, enhancing both reliability and energy efficiency. To address this limitation, this study presents a detailed classification of four different propagation environments using real traffic measurements and evaluates the performance of five advanced machine learning models: KNN, RF, Artificial Neural Networks (ANN), SVM, and AdaBoost (AB). Additionally, the Out-of-Bag (OOB) permutation feature importance method was employed to analyze the impact of environmental factors on classification performance. By integrating a holistic evaluation framework with state-of-the-art machine learning techniques, this study not only provides a comprehensive analysis of environmental classification in V2V networks but also offers valuable insights for optimizing communication systems in complex propagation environments.

The remainder of this paper is structured as follows: Section 2 provides a theoretical background on machine learning techniques used in this study. Section 3 details the measurement scenarios, dataset preparation, and model training process. Section 4 presents the results and discussion, highlighting key findings. Finally, Section 5 concludes the study and outlines future research directions.

2. Theory

2.1. Machine learning model

Machine learning is a fundamental artificial intelligence component that enables systems to learn and improve from data without explicit programming. It encompasses a variety of methods suited to different types of tasks, such as classification, regression, and clustering [24]. In this study, we examine methods such as SVM, AB, KNN, ANN, and ensemble techniques like RF, which have shown significant effectiveness in handling complex data patterns and making accurate predictions. This section provides an overview of these key machine-learning methods and their theoretical foundations.

2.1.1. Artificial neural networks

Inspired by the architecture and functioning of the human brain, these advanced computational models consist of interconnected processing elements, called neurons, that collectively address complex tasks like pattern recognition and optimization. Each artificial neuron calculates a weighted sum of inputs, processed through a non-linear activation function—commonly sigmoid or ReLU—to produce an output.

One of the foundational representations of an artificial neuron is the McCulloch-Pitts model:

$$y = \sigma\left(\sum_{j=1}^{n} w_j x_j + b\right) \tag{1}$$

where y represents the output, x_j denotes the input signals, w_j are the weights applied to each input, b is a bias term, and σ is an activation function [21]. A frequently used activation function is the sigmoid, defined as:

$$\sigma(z) = \frac{1}{1+e^{-z}} \tag{2}$$

By arranging these neurons into layers, these networks form intricate representations, especially in deep architectures with multiple hidden layers. Training is typically achieved using optimization algorithms like backpropagation to minimize errors and improve task performance [21].

2.1.2. Random forest

This ensemble learning method leverages the power of multiple decision trees to enhance prediction accuracy, especially in classification and regression tasks. By training on randomly selected subsets of the input features, the approach reduces correlation between individual trees, thereby minimizing overfitting. The generalization error PE^* depends on the strength of individual trees and the correlation between them. Strong trees with low inter-tree correlation ensure reduced errors across the forest. The expected generalization error is formalized using the margin function:

$$mg(X,Y) = \text{avg}_k\, I(h_k(X) = Y) - \max_{j \neq Y} \text{avg}_k\, I(h_k(X) = j) \tag{3}$$

where $I(\cdot)$ is an indicator function. This formulation highlights the confidence in classifications, where higher margins lead to lower errors [22,23].

2.1.3. Support vector machines

This supervised learning algorithm focuses on finding an optimal hyperplane that maximizes the margin between classes in a dataset. For linearly separable data, the optimization problem is defined as:

$$\min_{w,b} \frac{1}{2}\|w\|^2 \tag{3}$$

subject to:

$$y_i(w \cdot x_i + b) \geq 1 \quad \forall i \tag{4}$$

In cases where data is not linearly separable, slack variables ξ_i allow for misclassifications:

$$y_i(w \cdot x_i + b) \geq 1 - \xi_i, \quad \xi_i \geq 0 \tag{5}$$

This modified formulation balances margin maximization and classification error minimization [24,25].

2.1.4. AdaBoost

Classification accuracy is improved by iteratively combining weak classifiers into a strong one. During each iteration, the algorithm adjusts sample weights to focus on harder cases, with the weight of each classifier determined by:

$$\alpha_t = \frac{1}{2}\ln\left(\frac{1-\epsilon_t}{\epsilon_t}\right) \tag{6}$$

where ϵ_t represents the classification error of the weak learner at iteration t, and α_t denotes the weight assigned to the weak classifier, determined based on its error rate [26,27]. The weight update rule is:

$$w_{t+1}(i) = w_t(i)\exp(-\alpha_t y_i h_t(x_i)) \tag{7}$$

where $w_t(i)$ represents the weight of sample i at iteration t, $w_{t+1}(i)$ denotes the updated weight of sample i for the next iteration, y_i corresponds to the true class label of sample i, $h_t(x_i)$ represents the predicted class label for sample i from the weak classifier at iteration t, and α_t is the weight assigned to the weak classifier. By concentrating on misclassified samples, this iterative process effectively reduces training errors [26,27].

2.1.5. KNN

In multi-label classification tasks, KNN algorithm identifies the nearest instances in the training set and computes label frequencies [28]. Using the Maximum A Posteriori (MAP) principle, the labels are determined based on the neighbors' class distributions:

$$\hat{y}(l) = \arg\max_{b\in\{0,1\}} P(H_l^b | E_j^l) \tag{8}$$

Here, H_l^b represents the event that the instance belongs ($b = 1$) or does not belong ($b = 0$) to label l, and E_j^l indicates exactly j neighbors belonging to label l. This probabilistic framework allows KNN to handle

multi-label settings effectively by leveraging neighborhood information [28].

The selection of machine learning models in this study—KNN, RF, ANN, SVM, and AB—was guided by their proven effectiveness in handling structured, tabular data and their computational feasibility for real-time applications such as V2V communication. These models provide a balance between interpretability, prediction accuracy, and training/inference efficiency. Although recent advances in deep learning, including CNN, LSTM, and Transformer-based models, have shown strong performance in various classification tasks, they often require large datasets, high computational resources, and long training times, which may not be suitable for lightweight or embedded V2V systems ([6,7,11,12]). Furthermore, the complexity of deep architectures can limit their interpretability, which is essential for critical safety applications. Similarly, while ensemble models like XGBoost offer competitive accuracy, they demand more hyperparameter tuning and longer training cycles compared to the selected methods [13]. Hence, this study prioritizes models that strike an optimal balance between performance and practicality in resource-constrained vehicular environments.

2.1.6. Out-of-bag permutation feature importance (OOB-PFI)

OOB Permutation Feature Importance is a technique widely used in RF models to assess the relative importance of features. In this method, the OOB samples (data points not used in training a particular tree) are leveraged to evaluate model accuracy. The approach involves randomly permuting the values of a specific feature and measuring the decrease in model performance on the OOB samples. If permuting a feature results in a significant drop in accuracy, it indicates that the feature is crucial for the model's predictions. Conversely, if the impact is minimal, the feature is deemed less important. This method not only provides insights into feature relevance but also mitigates biases, especially for categorical variables with numerous categories, by ensuring that the importance score is not artificially inflated. Thus, OOB permutation offers a robust and data-efficient mechanism for feature selection and model interpretation without requiring additional cross-validation procedures [29].

2.2. Model evaluation metrics

The evaluation of machine learning models is crucial to assess their effectiveness and reliability, especially in supervised learning tasks. Various metrics are used to evaluate the performance of models based on their predictions. Commonly utilized metrics include accuracy, precision, recall, F1-score, and area under the receiver operating characteristic (ROC) curve. These metrics provide insights into different aspects of model performance, such as its ability to accurately classify data, its sensitivity to positive cases, and its capacity to minimize false positives. For instance, sensitivity (or recall) and specificity are often paired to better understand model effectiveness in imbalanced datasets, where accuracy alone can be misleading. Additionally, statistical tests like McNemar's test and the DeLong test are employed to compare the performance of different models in terms of metrics like sensitivity and AUC [30]. These evaluation techniques ensure that the selected models are robust and generalize well to unseen data.

Accuracy measures the proportion of correct predictions made by the model over the total number of predictions. It is particularly effective for balanced datasets but can be misleading for imbalanced datasets [31].

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{9}$$

where TP represents true positives, TN denotes true negatives, FP refers to false positives, and FN corresponds to false negatives.

Precision measures the proportion of correctly predicted positive instances out of all instances predicted as positive [31].

$$\text{Precision} = \frac{TP}{TP + FP} \tag{10}$$

Recall (also known as sensitivity) indicates the proportion of actual positive instances that are correctly identified by the model [31].

$$\text{Recall} = \frac{TP}{TP + FN} \tag{11}$$

The F1 score is the harmonic mean of Precision and Recall, providing a balance between the two, especially useful in cases where there is an uneven class distribution [31].

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{12}$$

AUC-ROC measures the overall ability of the model to distinguish between positive and negative classes. A higher AUC indicates better model performance [31].

$$AUC = \frac{S_p - S_n\left(\frac{n_p+1}{2}\right)}{n_p n_n} \tag{13}$$

where S_p represents the sum of the ranks of positive examples, n_p denotes the number of positive examples, and n_n corresponds to the number of negative examples.

In classification problems, the confusion matrix is a useful tool to evaluate the performance of a model. It is a table that summarizes the performance of a classification model by comparing the predicted values with the actual values [32]. The matrix consists of four components:

In a binary classification problem, True Positives (TP) refer to instances that are correctly predicted as positive, while True Negatives (TN) represent instances correctly predicted as negative. Conversely, False Positives (FP) occur when negative instances are incorrectly classified as positive (also known as Type I Error), and False Negatives (FN) arise when positive instances are mistakenly classified as negative (also known as Type II Error). The confusion matrix for binary classification is represented as follows:

$$\begin{bmatrix} \text{TP} & \text{FP} \\ \text{FN} & \text{TN} \end{bmatrix} \tag{14}$$

3. Measurement campaign and methodology

3.1. Measurement scenarios and setup

The dataset used in this study was collected over a three-month period between July and December 2023 in the provinces of Trabzon and Gümüşhane, Turkey. During this period, extensive drive tests were conducted in diverse environments—rural, highway, suburban, and urban—under varying weather and traffic conditions to ensure a representative sample. The environmental labels were assigned based on GPS-tagged location data and cross-verified with Google Maps and satellite imagery to minimize labeling uncertainty. Manual verification was also performed to ensure that borderline segments (e.g., transitions between suburban and urban areas) were accurately annotated. As the labeling process was both rule-based and operator-verified, the uncertainty in class definitions is considered minimal. This approach contributes to the reliability and reproducibility of the dataset. Each environment presents unique characteristics that influence the classification performance. A detailed analysis of the measurement environments highlights their distinct propagation characteristics, which are primarily shaped by the nature and distribution of the obstacles. Urban environments are defined by the presence of tall buildings along measurement routes, dense yet slow-moving traffic (typically 20–40 km/h), frequent



Fig. 1. Measurement route for highway environment in Trabzon and Gümüşhane.



Fig. 2. Measurement route for urban environment in Trabzon and Gümüşhane.

obstructions between the transmitter (Tx) and receiver (Rx) caused by vehicles, as well as infrastructure elements such as traffic signals, streetlights, and pedestrian crossings. These areas generally consist of one or more traffic lanes. Suburban environments, located at the periphery of city centers, exhibit lower building and vehicle densities, moderate traffic flow (40–60 km/h)—slower than highways but faster than urban areas—and roadside vegetation, including trees and medians. Highways are characterized by high-speed traffic (averaging 90–110 km/h) owing to elevated vehicle density, multi-lane roads, and the absence of significant obstacles such as buildings and trees. Rural environments are situated at considerable distances from city centers and consist mainly of narrow single- or dual-lane roads with horizontal and sloped topographies, minimal vehicle density leading to low traffic volumes, and an abundance of natural obstacles such as trees and mountains, which set them apart from other environments. Field measurements were conducted over multiple days to ensure data diversity, capturing both clear and adverse weather conditions, including sun and snowfall. These variations are critical in understanding the impact of environmental factors on V2V communication reliability. The routes of the measurement environments are shown in Figs. 1–4. Measurements for highway, suburban, and urban environments were conducted in both Trabzon and Gümüşhane, whereas rural environment measurements were made exclusively in Gümüşhane.

Fig. 5 presents a block diagram of the measurement system used for the data collection. The system incorporates two laptops connected to the DSRC OBU units installed in both Tx and Rx vehicles. These vehicles are equipped with DSRC + GPS modules that enable real-time data collection. For the data acquisition and processing phases, each vehicle was fitted with a DSRC OBU device, the technical specifications of which are listed in Table 2. In this study, two passenger cars were utilized for experimental measurements: a 2022 Škoda Fabia transmitter and a 2013 Ford Fiesta receiver. Additionally, a 12V/220V converter and in-car camera were installed inside the vehicles. The laptops were connected to the DSRC OBU device via an Ethernet port, allowing for the control of data transmission. Fig. 6 provides an interior view of the transmitter and receiver vehicles, showing the MK5 OBU kit antenna, laptop, and power supply components. The devices were powered by multiple converters to ensure stable operation throughout the experiment.

3.2. Data processing

Fig. 7 illustrates the input parameters used in various machine learning models for environmental classification. To predict rural, highway, suburban, and urban propagation environments, the key input features include path loss, distance, obstacle types weighted by path loss and distance, modulation schemes weighted by path loss, and weather conditions weighted by path loss. Propagation environmental classification plays a crucial role in V2V communication by enabling dynamic adjustments to the transmission power of the communication devices. For example, in suburban areas, where increased vehicle density may lead to communication disruptions, increasing the transmission power can help maintain connectivity. Conversely, in rural environments, where there are no obstacles between two autonomous vehicles, reducing the

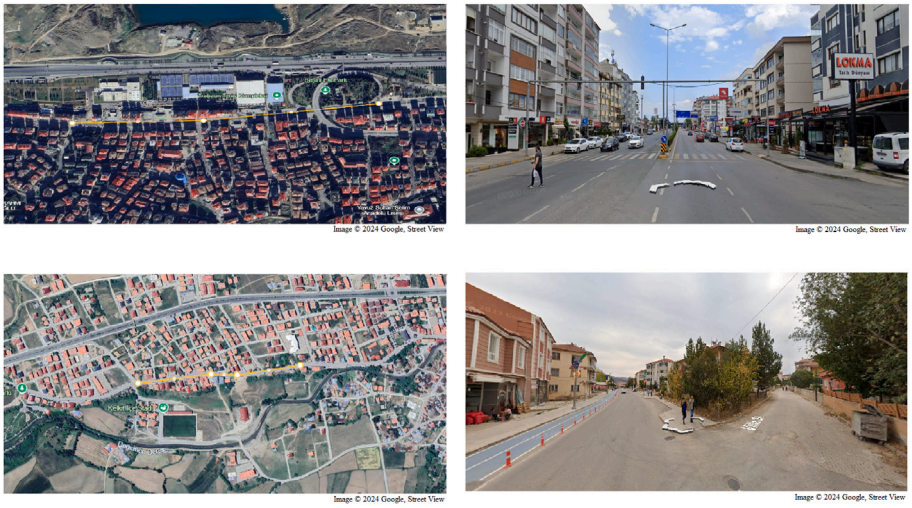


Fig. 3. Measurement route for suburban environment in Trabzon and Gümüşhane.



Fig. 4. Measurement route for rural environment in Gümüşhane.

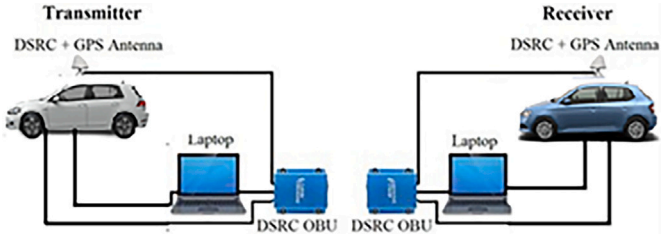


Fig. 5. Measurement System Block Diagram [33].

Table 2
Technical specifications of DSRC OBU device [34].

Parameter	Value
Standard	IEEE 802.11p
Frequency band	5.9 GHz
Data rate	3–27 Mbps
Tx power	22 dBm
Antenna gain	5 dBi
Antenna heights	Vehicles = 1.46 m (Tx and Rx)
Rx sensitivity	−99 dBm at 3 Mbps

transmission power enhances energy efficiency, allowing for prolonged operation with limited power resources. Real-time power optimization is particularly important in modern V2V systems because it ensures that vehicles maximize the efficiency of their available power while maintaining reliable communication.

As illustrated in Fig. 7, the system utilizes a combination of physical and contextual input features to accurately classify the propagation environment. Distance, ranging from 10 to 300 m, captures the spatial separation between vehicles and is a fundamental factor influencing path loss. The number of obstacles, varying from 0 to 5 in the dataset, reflects line-of-sight obstruction, which is particularly critical in urban and suburban scenarios. Obstacle type (e.g., static buildings or mobile vehicles) is encoded categorically to represent different attenuation characteristics. Weather conditions (e.g., clear, rainy, foggy) are also included due to their influence on signal degradation, especially at higher frequencies. Additionally, modulation type and communication channel parameters are integrated as auxiliary features to represent the physical layer's adaptation to changing environments. Each of these features contributes to a more reliable and rapid classification process, enabling the system to better adapt to dynamic V2V conditions and support real-time decision-making.

All simulations and analyses were performed using MATLAB R2024b on a computer equipped with a 64-bit Windows 10 Pro operating system, an Intel® Core™ i5-9300H CPU with 2.40 GHz, 16.0 GB RAM, and an x64-based processor. To mitigate the effects of varying data scales, a normalization process, as outlined in Wu et al. [35], was applied. This ensures that all features contribute equally to the model, enhancing its stability and efficiency and resulting in more accurate outcomes and shorter convergence times during training. The formula used for normalization is as follows:

$$x_n = \frac{2(x - x_{\min})}{x_{\max} - x_{\min}} - 1$$

where, x_n represents the normalized value, while x denotes the original value. The values $x_{\min}$ and $x_{\max}$ refer to the minimum and maximum values within the dataset, respectively. To ensure that all input features contribute equally to the classification models, a normalization process was applied, scaling all values between -1 and 1. This transformation mitigates the effect of varying data magnitudes, improving model convergence and stability, particularly in algorithms such as SVM and ANN, which are sensitive to feature scaling. This normalization technique also facilitates a fast and effective learning performance by addressing issues related to varying data scales. After completing the data cleaning and preprocessing steps, 80% of the dataset was allocated as the training set

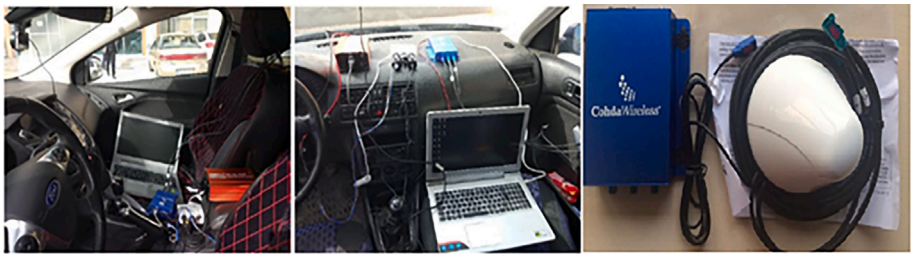


Fig. 6. Interior view of the Tx and Rx vehicle and MK5 OBU Kit [34].

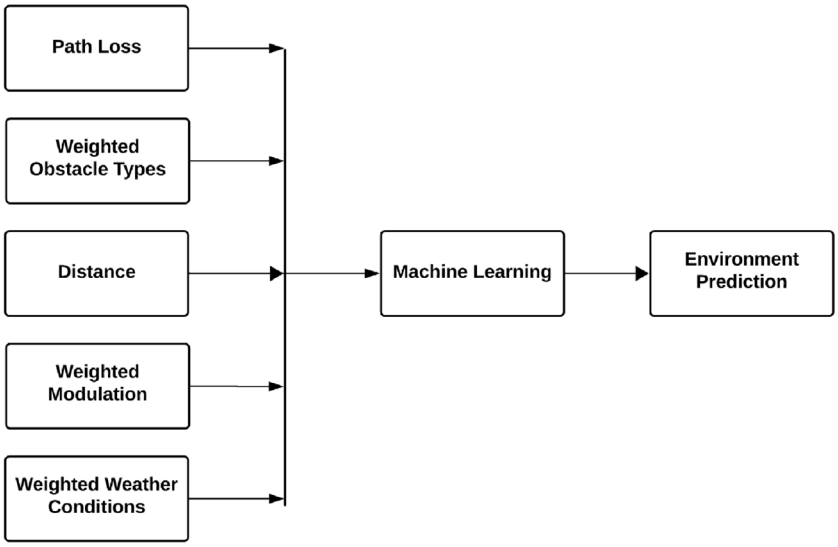


Fig. 7. Input parameters used in environment propagation classification prediction with machine learning.

Table 3
Hyperparameter optimization for classification methods.

Model	Parameter	Tested range	Optimized value	Optimization method
SVM	Kernel Function	RBF, Linear, Polynomial	RBF	Grid Search
	C (Regularization)	0.1–0.5–1–5–10	0.5	Grid Search
	Gamma	0.001–0.01–0.1–1	0.01	Grid Search
	Kernel Scale	0.1–1–5–10	1	Grid Search
KNN	Number of Neighbors	1–3–5–7–9	5	Random Search
	Weight Function	Uniform, Distance	Distance	Random Search
	Algorithm	Auto, BallTree, KDTree	Auto	Random Search
RF	Number of Estimators	50-100-200-300	200	Bayesian Optimization
	Max Features	Auto, sqrt, log2	sqrt	Bayesian Optimization
	Max Depth	1–5–10	10	Bayesian Optimization
	Min Samples Split	2–5–10	2	Bayesian Optimization
	Criterion	Gini, Entropy	Gini	Bayesian Optimization
ANN	Number of Layers	3	3	Grid Search
	Neurons in 1st Layer	0–10–50–100	10	Grid Search
	Neurons in 2nd Layer	0–10–50–100	10	Grid Search
	Neurons in 3rd Layer	0–10–50–100	10	Grid Search
	Activation Function	ReLU, Sigmoid, Tanh	ReLU	Grid Search
	Epochs	500–1000–1500–2000	1500	Grid Search
	Batch Size	32–64–128	64	Grid Search
AB	Base Estimator	Decision Tree, RF	Decision Tree	Genetic Algorithm
	Number of Estimators	50–100–150–200	100	Genetic Algorithm
	Learning Rate	0.01–0.1–0.5–1	0.1	Genetic Algorithm

and the remaining 20% was reserved for model validation as the test set. To evaluate the generalization capability of the model accurately, a five-fold cross-validation approach was employed.

Following the parameter adjustments, Table 3 highlights the detailed hyperparameter optimization process aimed at enhancing the classification performance of machine-learning models for V2V propagation environments. For the RF model, 200 estimators were selected to strike a balance between bias and variance. The “sqrt” setting for maximum features facilitated efficient computation and improved accuracy by selecting a subset of important features. The SVM model employed

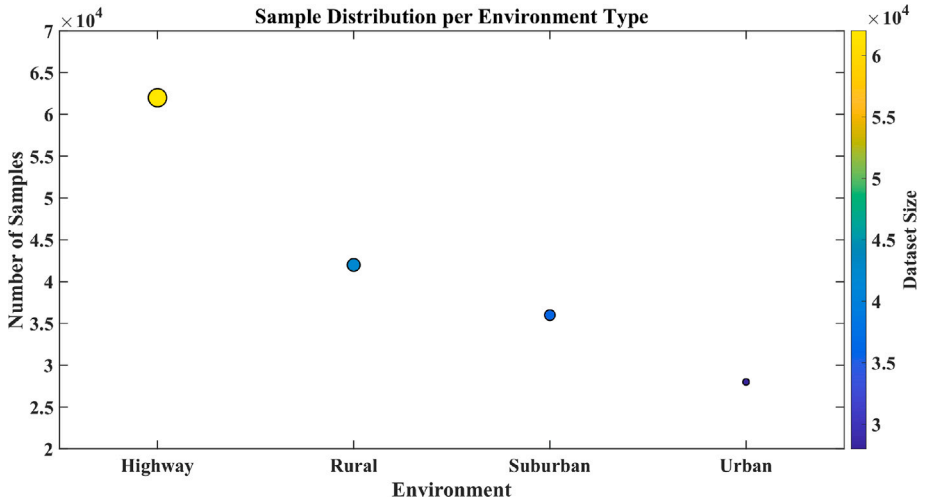


Fig. 8. Number of samples per environment.

the Radial Basis Function (RBF) kernel to effectively capture nonlinear relationships, with a regularization parameter C of 0.5, and a gamma value of 0.01, to promote robust generalization and minimize overfitting. For KNN, five neighbors were used with distance-based weighting, allowing the model to adapt effectively to variations in feature distributions and enhance its classification ability across diverse environments. The ANN model, configured with three layers to balance model complexity and computational efficiency while effectively capturing nonlinear relationships and ReLU activation, demonstrated its capacity to learn complex relationships, supported by 1500 epochs and a batch size of 64 to ensure stable convergence without overfitting. Finally, AB utilized decision trees as the base estimator, with 100 estimators and a learning rate of 0.1, providing a balanced ensemble capable of handling diverse and complex input features. This systematic approach to hyperparameter tuning not only emphasizes accuracy but also improves the generalizability of the models, making them well-suited for the reliable classification of V2V communication environments in intelligent transportation systems.

Additionally, Fig. 8 shows the number of samples for each environment type—highway, rural, urban, and suburban—in the dataset, which consists of a total of 111,520 observations.

4. Results and discussion

In this section, we present the classification results of various machine-learning models across different propagation environments. The main goal was to evaluate how effectively each model distinguishes between rural, highway, suburban, and urban areas, each with unique propagation characteristics. To measure model performance, we used accuracy, precision, recall, F1 score, and AUC-ROC metrics. We also employed the OOB permutation feature importance technique to assess the relative contribution of each input feature. This method provides an unbiased estimate of feature relevance, helping to identify the most influential factors for accurate classification. The results in Table 4 indicate that the KNN model consistently outperforms the others. It achieved the highest scores across all metrics, with an accuracy of 0.9984, precision of 0.9983, recall of 0.9984, F1 score of 0.9984, and AUC-ROC of 0.9997. While RF and ANN slightly surpassed KNN in AUC-ROC, KNN's consistent performance across all other metrics highlights its overall superiority. The outstanding performance of KNN (99.84% accuracy) can be attributed to its instance-based learning strategy, which preserves local patterns in the feature space. Since our dataset includes well-separated environmental classes and weighted features such as path loss and obstacle types, KNN is able to capture nuanced propagation differences more effectively than parametric models. Its non-parametric nature allows flexible decision boundaries that adapt to subtle environmental changes, especially in transition zones. This local adaptability is particularly advantageous when dealing with high-dimensional but structured tabular data, as observed in our V2V dataset. RF ranked second overall, while SVM delivered the weakest performance, with an accuracy of 0.8479 and F1 scores of 0.8498 and 0.9204. To assess whether the accuracy differences among the top models are statistically significant, we conducted a paired *t*-test and a Wilcoxon signed-rank test based on the 5-fold cross-validation results of KNN and RF. The paired *t*-test yielded a p-value of 0.00084, indicating a statistically significant performance difference favoring KNN. The Wilcoxon test also supported this trend, albeit with a p-value of 0.0625. These tests confirm that the observed accuracy difference, although small in magnitude (0.02%), is not due to random variation. To rule out the possibility of overfitting, we ensured a strict separation between training and test datasets and applied 5-fold cross-validation. The use of a large and balanced dataset, containing 111,520 samples equally distributed among the four classes, further reduces this risk. The relatively poor performance of SVM may stem from its sensitivity to feature scaling and the non-linear nature of obstacle-related features, which may not align well with its kernel assumptions. Overall, the differences observed across models reflect their underlying learning mechanisms rather than overfitting tendencies. Notably, KNN's ability to handle high-dimensional and structured data makes it especially well-suited to the task.

Table 4
Output performance of methods in the classification problem (Test).

Model/method	Accuracy	Precision	Recall	F-1 score	AUC-ROC
KNN	0.9984	0.9983	0.9984	0.9984	0.9997
RF	0.9982	0.9980	0.9981	0.9980	1.0000
AB	0.9977	0.9975	0.9975	0.9975	0.9970
ANN	0.9939	0.9936	0.9936	0.9936	1.0000
SVM	0.8479	0.8406	0.8731	0.8498	0.9204

In real-time classification systems, not only classification accuracy but also computational efficiency plays a critical role in practical deployments. To this end, Table 5 presents the training time and per-sample inference time of each evaluated machine learning model. These metrics are essential for assessing the feasibility of deploying such models in resource-constrained vehicle-to-vehicle (V2V) communication systems. The KNN model, in particular, demonstrates the lowest training and inference times among the tested models, in addition to achieving the highest classification performance (see Table 4). While models

Table 5
Computational efficiency of the models.

Model/method	Training time(s)	Inference time per sample(ms)
KNN	1.8	0.07
RF	5.6	0.11
AB	7.1	0.16
ANN	14.9	0.24
SVM	22.3	0.28

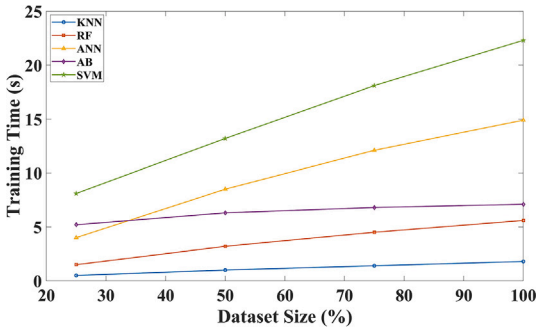


Fig. 9. Scalability of training time with dataset size.

such as ANN and SVM exhibit strong classification capabilities, their relatively high computational costs may hinder their application in latency-sensitive and real-time environments. These findings emphasize that model selection for V2V systems should not solely be based on predictive accuracy but must also account for system latency and processing overhead to ensure timely and efficient environmental classification in dynamic traffic scenarios.

To examine the scalability of the models, training time was measured for different dataset sizes (25%, 50%, 75%, and 100%). As shown in Fig. 9, KNN and RF demonstrate relatively low and linear growth in training time, making them suitable for large-scale deployment. In contrast, SVM and ANN show steeper increases, indicating higher computational demands as the dataset expands.

Fig. 10(a)–(e) show the confusion matrices illustrating the classification performance of the evaluated methods. Class 1, Class 2, Class 3, and Class 4 correspond to rural, highway, suburban, and urban environments, respectively. Upon examining the confusion matrix of the best-performing KNN algorithm, it is observed that it misclassified a total of 31 instances while correctly classifying 22,261 instances. The most significant misclassification occurred when 13 instances were predicted as Class 2 (highway) instead of their actual Class 4 (urban). The weakest performance was exhibited by the SVM model, as seen in Fig. 10(e), where the most significant classification error involved 1351 instances in which the actual class was Class 4 (urban), but the model predicted them as Class 3 (suburban). This error was attributed to the similarity of input parameters between urban and suburban environments. The confusion matrix of the second-best performing method, RF, shown in Fig. 10(d), indicates that it misclassified a total of 50 instances while correctly classifying 22,254 instances. In the confusion matrix of the ANN model, it is noted that the model struggled to distinguish between certain classes, particularly Classes 2 and 4, indicating its limited ability to learn complex decision boundaries. In contrast, the KNN model showed some false positives in Class 4 but generally maintained a strong discriminative ability across classes. Overall, these findings highlight the superiority of KNN in propagation environment classification, followed by RF and ANN, while SVM exhibited the weakest performance due to its difficulties in differentiating between urban and suburban environments.

The ROC curves in Fig. 11(a)–(e) provide a comprehensive evaluation of the classification performance of various machine learning models for environment classification in V2V communication systems. In particular, the AB model (Fig. 11(a)) and ANN (Fig. 11(b)) show exceptional AUC values close to 1.0 in all classes, highlighting their robustness in differentiating complex environments such as urban and suburban areas. This superior performance can be attributed to the models' ability to capture intricate nonlinear relationships between input parameters like path loss, distance, and obstacle types weighted by their impact, as discussed in Section 4 of the study. The RF model (Fig. 11(d)) demonstrates a similarly high AUC, using its ensemble nature to reduce variance and improve prediction accuracy, particularly under dynamic environmental conditions. In contrast, while the SVM model (Fig. 11(e)) exhibits relatively lower AUC values, especially for urban and suburban classifications, its sensitivity to high-dimensional feature spaces allows it to maintain robust classification in scenarios with overlapping feature distributions. Finally, the KNN model (Fig. 11(c)) achieves near-perfect classification accuracy by effectively using neighborhood-based decision boundaries, which is particularly beneficial in environments with different propagation characteristics. These results underscore the efficacy of ensemble and nonlinear classifiers in optimizing real-time V2V communication by dynamically adjusting transmission power based on environmental context, thus enhancing communication reliability and efficiency in intelligent transportation systems.

Then OOB Permutation Feature Importance method was utilized to determine the relative impact of different input parameters on environmental classification, as seen in Table 6, which provides a quantitative comparison of feature contributions to model predictions. This method employs OOB samples inside a random forest framework to measure feature relevance by methodically permuting the values of each feature and evaluating the resulting effect on model accuracy. A significant decline in accuracy upon permutation indicates that this corresponding property is essential for precise predictions. The examination concluded that factors such as distance and weighted obstructions, particularly buildings and mountains, had the most significant impact, highlighting their crucial role in accurate environmental identification. This method offers an extensive and objective evaluation of the relevance of features, thereby improving the adaptability and prediction performance of the model in complex communication scenarios.

Finally, the correlation analysis revealed a strong positive relationship (0.6756) between distance and path loss, confirming that signal strength decreases with increasing distance in V2V communication systems. This trend is clearly depicted in the correlation matrix (Fig. 13). High correlations among weighted environmental features—such as buildings and cars (0.9275)—reflect their frequent co-occurrence in dense urban areas, where they significantly impact signal propagation. Similarly, modulation schemes like QPSK and BPSK show a strong correlation with weighted building features (0.9202), indicating their sensitivity to complex environments and higher attenuation. Weather-related features, such as snowfall in mountainous regions, also exhibit a moderate correlation (0.638), suggesting that adverse weather further impairs communication reliability. These results highlight the value of incorporating weighted environmental and contextual features into machine learning models. They also support the study's conclusion that environment-aware signal adjustments can improve both reliability and efficiency in real-world V2V communication scenarios. To further validate the importance of individual obstacle-related features, we performed an ablation analysis by selectively removing each weighted obstacle (building, mountain, vehicle, and tree) from the input set. As shown in Fig. 12, removing the building feature led to the highest accuracy drop (2.81%), followed by mountain (2.13%), vehicle (1.96%), and tree (1.25%). This experiment confirms that the classification model heavily relies on obstacle information, particularly the presence of buildings and mountainous terrain.

Previous studies in the literature have generally focused on various classification tasks in vehicular communication, such as driving environment classification with traditional machine learning models achieving up to 93% accuracy [8] and vehicle classification using deep

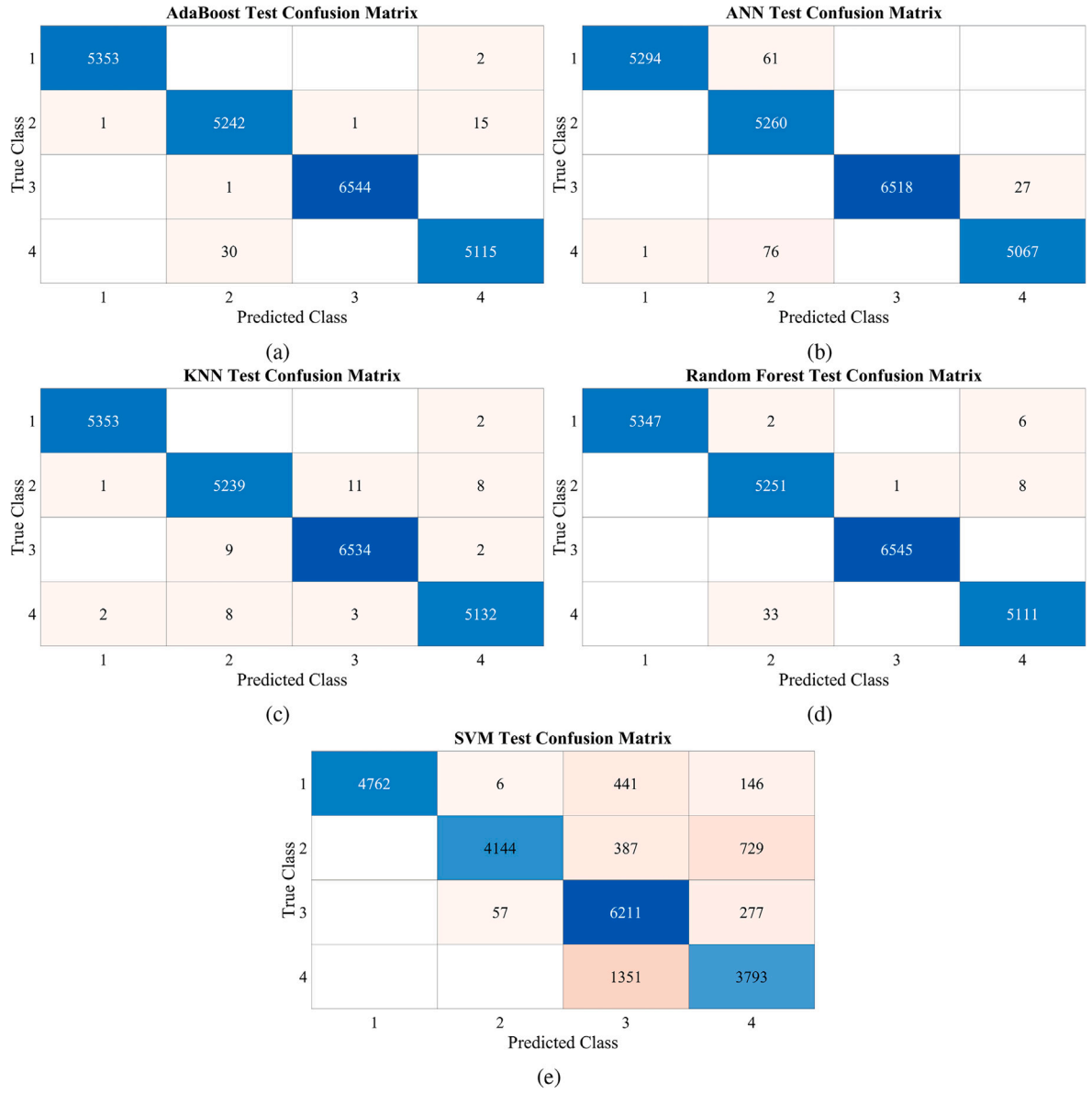
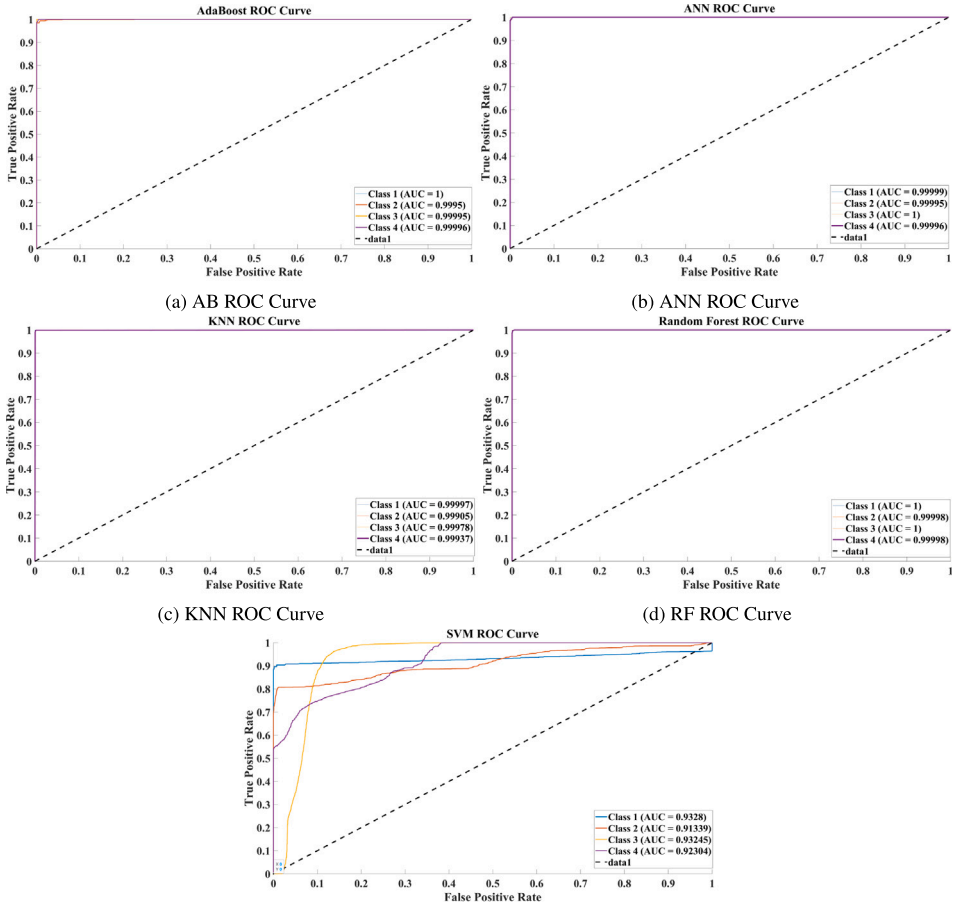


Fig. 10. Confusion matrices for environment classification problem using different machine learning models: (a) AB, (b) ANN, (c) KNN, (d) RF, and (e) SVM.

learning methods with an F1-score of 99.56% [6]. Additionally, satellite image-based path loss prediction studies have reached 89.41% accuracy [36]. However, these studies either focus on limited environments and specific classification tasks or do not incorporate real-world measurement data for comprehensive V2V communication analysis. Unlike these studies, our research provides a holistic approach by integrating multi-environment classification using experimental measurements and machine learning models tailored for V2V propagation scenarios. The significantly higher classification accuracy achieved in this study (99.84% with KNN) demonstrates the effectiveness of leveraging weighted environmental parameters. Furthermore, our findings highlight the necessity of environment-aware adaptive transmission strategies in V2V communication, ensuring enhanced reliability and efficiency. To sum up, these results emphasize the importance of integrating real-time environmental classification into intelligent transportation systems to optimize communication performance under diverse conditions.

As seen in Table 7, most of the prior studies have focused on either limited environments or specific tasks such as vehicle type [6,13], anomaly detection [14], or modulation classification [18]. Deep learning approaches like CNNs [11,12] and LSTMs [7] provide high accuracy but typically require large-scale data, substantial computational resources, and lack interpretability. Moreover, studies such as Julio-Rodríguez et al. [8] and Wang et al. [7] are restricted to urban or highway environments only, limiting their generalization to diverse scenarios. In contrast, our study utilizes real-world field data collected from four distinct propagation environments and systematically evaluates five classical machine learning models in terms of both predictive accuracy and computational efficiency. With an achieved accuracy of 99.84% using KNN, our model not only outperforms the state-of-the-art methods but also offers a lightweight and interpretable solution suitable for real-time implementation in V2V systems.

To demonstrate a practical application of our classification system, consider an adaptive power control algorithm that dynamically



(a) AB ROC Curve (b) ANN ROC Curve

(c) KNN ROC Curve (d) RF ROC Curve

(e) SVM ROC Curve

Fig. 11. ROC curves for different machine learning models.

Table 6
Effect of input parameters on the output parameter (Environment).

Input parameter	Effect on output parameter (Environment)
Distance	1.5853 (%9.96)
Path loss	1.1947 (%7.51)
Weighted obstacle (vehicle)	1.6437 (%10.33)
Weighted obstacle (tree)	1.0640 (%6.69)
Weighted obstacle (building)	3.1970 (%20.09)
Weighted obstacle (mountain)	2.2501 (%14.14)
Weighted modulation (QPSK)	1.1150 (%7.01)
Weighted modulation (BPSK)	1.5222 (%9.56)
Weighted weather (snowy)	1.1996 (%7.54)
Weighted weather (sunny)	1.1411 (%7.17)

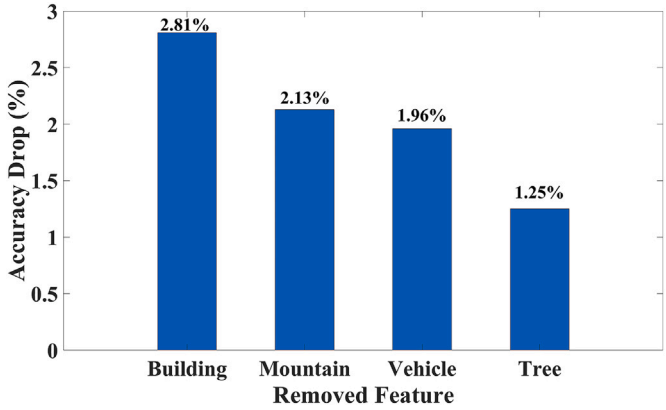


Fig. 12. Accuracy degradation after removing each weighted obstacle feature (building, mountain, vehicle, tree).

adjusts transmission power based on the classified environment. For example, when the system detects an urban or suburban environment—typically characterized by high obstacle density and frequent non-line-of-sight conditions—the transmission power can be increased within regulatory limits to overcome propagation loss and enhance link reliability. Conversely, in rural or highway environments, where line-of-sight conditions are dominant and interference is minimal, the transmission power can be reduced to conserve energy. This environment-aware power control strategy minimizes unnecessary energy consumption while maintaining reliable communication, particularly in dense vehicular networks. Such decisions can be implemented through DSRC or C-V2X stack layers by integrating the classification output as an input trigger to PHY/MAC layer parameter tuning modules.

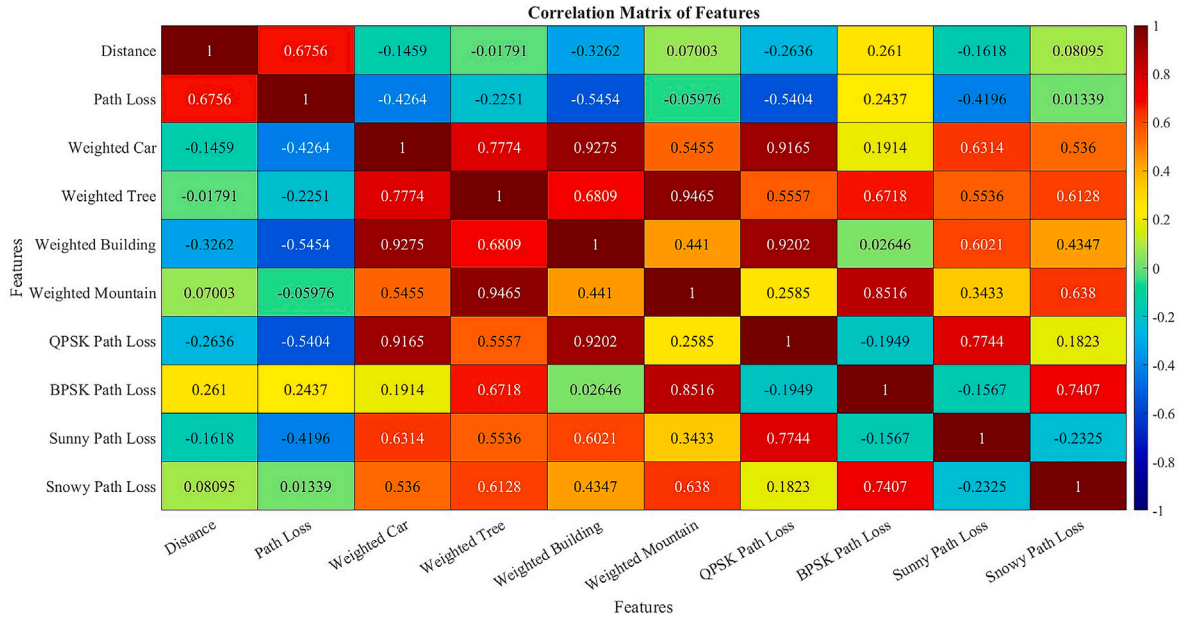


Fig. 13. Correlation matrix of features.

Table 7
Comparative analysis of related classification-based studies in the literature.

Reference	Environment	Method	Performance (Accuracy)
[6]	Various	CNN (ResNet)	99.68%
[7]	Highway	LSTM, Transformer	93%
[8]	Urban, Suburban	SVM, RF	93%
[11]	Satellite	CNN, Deep Learning	88%
[12]	Various	CNN, Hybrid Regression	88%
[13]	Various	XGBoost	98.81%
[14]	Ad-Hoc Network	BiGAN	92.15%
[18]	Wireless Channels	Semantic Learning	92–97%
This Study	Rural, Highway, Urban, Suburban	KNN, RF, ANN, SVM, AB	99.84%

5. Conclusion

In this study, the environment classification performance of well-known machine learning algorithms such as SVM, KNN, RF, ANN, and AB in V2V communication systems was investigated, leveraging real traffic measurements collected from different environments, including rural, highway, urban, and suburban areas. The results demonstrate that KNN, RF, and ANN models achieve superior classification performance, particularly in distinguishing complex urban and suburban environments where propagation characteristics are heavily influenced by obstacles such as buildings and vehicles. The KNN model, in particular, with an accuracy of 99.84%, and AUC-ROC of 99.97%, outperformed other models, proving its robustness in handling high-dimensional feature spaces and varied environmental contexts. The use of weighted input features, such as path loss, obstacle types, modulation schemes, and weather conditions, was found to significantly enhance classification accuracy; besides, the weighted building obstacle feature contributed the most (20.09%) to the overall classification success. Furthermore, employing the OOB Permutation Feature Importance method allowed to identify the critical impact of various input features on model performance, enabling a deeper understanding of their influence on V2V communication reliability. Although the primary objective was environment classification, the proposed system indirectly enhances vehicular safety by enabling real-time and accurate situational awareness. For instance, the KNN model's inference time of 0.07 ms facilitates ultra-fast decision-making, which can support safety-critical applications such as adaptive transmission power control, dynamic resource allocation, and early collision avoidance mechanisms. By allowing the network to recognize and respond to high-risk propagation conditions promptly, the proposed model lays a foundation for safer and more resilient V2V communication systems. Despite these promising results, the current study has certain limitations. The models rely on structured/tabular input features rather than real-time sensor or image data, which may limit adaptability in highly dynamic scenarios. Moreover, no external validation across different geographical locations or hardware platforms was performed, which could affect generalizability. These constraints should be addressed in future studies to ensure broader applicability. Future implementations may integrate the proposed classification system into DSRC or C-V2X middleware layers to enable real-time adjustments in transmission power and modulation schemes. Regulatory constraints on power control and system-level privacy policies should also be considered to ensure safe and compliant deployment. Conclusively, these findings indicate that dynamic adjustments of transmission power, based on real-time environmental classification, can optimize communication efficiency and reduce power consumption, particularly in autonomous vehicle networks. Notably, the RF and ANN models, with AUC values close to 1.0, exhibited strong generalization capabilities, particularly in environments characterized by varying obstacle densities and traffic patterns. Therefore, these insights underscore the potential of integrating machine learning-driven environment classification into intelligent transportation systems and autonomous vehicles, enabling adaptive communication strategies that enhance safety, reduce latency, and improve the overall efficiency of V2V communication networks. In addition, modern lightweight models such as MobileNet, TinyML, LightGBM, or CatBoost could be explored in future work to assess their effectiveness in low-latency V2V applications under dynamic conditions. In the future, to enable self-optimizing V2V communication networks, these models can be further enhanced by integrating real-time data streams and creating hybrid models that combine deep learning with conventional

machine learning techniques for even more accurate classification in extremely dynamic V2V environments.

CRediT authorship contribution statement

Nuğman Sağır: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Project administration, Methodology, Formal analysis, Data curation, Conceptualization. **Zeynep Hasırcı Tuğcu:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Karedal J, Czink N, Paier A, Tufvesson F, Molisch AF. Path loss modeling for vehicle-to-vehicle communications. IEEE Trans Veh Technol Jan 2011;60(1):323–8.

[2] Lyu F, Zhu H, Cheng N, Zhou H, Xu W, Li M, Shen X. Characterizing urban vehicle-to-vehicle communications for reliable safety applications. IEEE Trans Intell Transp Syst Jun 2020;21(6):2586–97.

[3] Yang X, Liu J, Zhao F, Vaidya NH. A vehicle-to-vehicle communication protocol for cooperative collision warning. In: Proc. IEEE International Conference on Mobile and Ubiquitous Systems: Networking and Services; 2004. p. 114–23.

[4] Ameen HA, Mahamad AK, Saon S, Nor DM, Ghazi K. A review on vehicle-to-vehicle communication system applications. Indones J Electr Eng Comput Sci Apr 2020;18(1):188–98.

[5] Zeadally S, Guerrero-Ibanez J, Contreras-Castillo J. A tutorial survey on vehicle-to-vehicle communications. Telecommun Syst Dec 2019;73(3):469–89.

[6] Butt MA, Awais M, Khan S, Rasheed A. Convolutional neural network-based vehicle classification in adverse illuminous conditions for intelligent transportation systems. Complexity 2021;2021:1–10. https://doi.org/10.1155/2021/9948937

[7] Wang S, Li C, Yang B. Accurate classification and prediction of remote vehicle position classes using V2V communication and deep learning. IEEE Trans on Vehicular Technology 2024;73(5):765–74. https://doi.org/10.1109/TVT.2024.3211110

[8] Julio-Rodríguez JC, Pérez A, García F. Uses of artificial intelligence in autonomous driving and V2X communication. Appl Sci 2022;12(5578):1–18. https://doi.org/10.3390/app12055578

[9] Boskovich S, Barth M. Vehicular network rerouting autonomy with a V2V, I2V, and V2I communication matrix classification. In: Proceedings of the 16th International IEEE Annual Conference on Intelligent Transportation Systems (ITSC 2013), The Hague. Netherlands. Oct 2013. p. 172–7. https://doi.org/10.1109/ITSC.2013.6728343

[10] Neunerdt M, Engels A, Mathar R. Land use classification as a key component for path loss prediction in rural areas. In: Proc. IEEE ISWCS. Aachen, Germany; 2010. p. 660–6. https://doi.org/10.1109/ISWCS.2010.5666654

[11] Ates HF, Hashir SM, Baykas T, Gunturk BK. Path loss exponent and shadowing factor prediction from satellite images using deep learning. IEEE Access 2019;7:101366–76. https://doi.org/10.1109/ACCESS.2019.2931072

[12] Sani US, Lai DTC, Malik OA. A hybrid combination of a convolutional neural network with a regression model for path loss prediction using tiles of 2D satellite images. In: Proc. 8th Int. Conf. Intelligent and Advanced Systems (ICIAS). Kuala Lumpur, Malaysia; 2020. p. 573–9. https://doi.org/10.1109/ICIAS49414.2020.9642585

[13] Pemila M, Pongiannan RK, Narayanamoorthi R, Sweelem EA, Hendawi E, El-Sebah MIA. Real-time classification of vehicles using machine learning algorithm on the extensive dataset. IEEE Access 2024;12:1–15. https://doi.org/10.1109/ACCESS.2024.3417436

[14] Khalil A, Farman H, Nasralla MM, Jan B, Ahmad J. Artificial intelligence-based intrusion detection system for V2V communication in vehicular adhoc networks. Ain Shams Eng J 2024;15. https://doi.org/10.1016/j.asej.2023.102616

[15] Abdellatif AA, Chiasserini CF, Malandrino F. Active learning-based classification in automated connected vehicles. IEEE Trans Intell Transp Syst Feb 2020;21(2):679–90. https://doi.org/10.1109/TITS.2020.2973020

[16] Barbieri L, Savazzi S, Nicoli M. Decentralized federated learning for road user classification in enhanced V2X networks. In: 2021 IEEE International Conference on Communications (ICC). Montreal, QC, Canada; 2021. p. 1–6. https://doi.org/10.1109/ICC42927.2021.9500867

[17] Tahmasbi-Sarvestani A, Mahjoub HN, Fallah YP, Moradi-Pari E, Abuchaar O. Implementation and evaluation of a cooperative vehicle-to-pedestrian safety application. IEEE Intelligent Transportation Systems Magazine 2017;9(4):6–17, Winter https://doi.org/10.1109/MITS.2017.2743201

[18] Kutay E, Yener A. Classification-oriented semantic wireless communications. [arXiv preprint] arXiv:2401.18069. 2024.

[19] Xing H, Zhang X, Chang S, Ren J, Zhang Z, Xu J, Cui S. Joint signal detection and automatic modulation classification via deep learning. [arXiv preprint] arXiv:2405.00736. 2024.

[20] Tangade S, Manvi SS, Hassan S. A deep learning-based driver classification and trust computation in VANETs. IEEE Access 2024;12:5634–45. https://doi.org/10.1109/ACCESS.2024.1234567

[21] Jain AK, Mao J, Mohiuddin KM. Artificial neural networks: A tutorial. Computer 1996;29(3):31–44. https://doi.org/10.1109/2.485891

[22] Breiman L. Random forests. Mach Learn 2001;45:5–32. https://doi.org/10.1023/A:1010933404324

[23] Sağır N, Hasırcı Tuğcu Z. A comparative analysis of machine learning models based on weighted input parameters for V2V path loss prediction in highway, rural, suburban, and urban environments. Comput Electr Eng 2025;128:110722. https://doi.org/10.1016/j.compeleceng.2025.110722

[24] Wang H, Ma C, Zhou L. A brief review of machine learning and its application. In: 2009 International Conference on Information Engineering and Computer Science; 2009. https://doi.org/10.1109/ICIECS.2009.5362979

[25] Alper A, Sağır N, Tuğcu ZH. Obstacle-aware machine learning classification in V2V communication based on path loss patterns. In: 2025 9th International Artificial Intelligence and Data Processing Symposium (IDAP). Malatya, Türkiye; 2025. p. 1–6. https://doi.org/10.1109/IDAP68205.2025.11222174

[26] Rätsch G, Onoda T, Müller KR. Soft margins for AdaBoost. Mach Learn 2001;42:287–320. https://doi.org/10.1023/A:1007618119488

[27] Sağir N, Tuğcu ZH. Comparative analysis of ANN, SVR, and AdaBoost models for V2V path loss prediction in urban and suburban environments. In: 2024 8th International Artificial Intelligence and Data Processing Symposium (IDAP). Malatya, Turkiye; 2024. p. 1–5. https://doi.org/10.1109/IDAP64064.2024.10711139

[28] Zhang M-L, Zhou Z-H. ML-KNN: A lazy learning approach to multi-label learning. Pattern Recognit 2007;40:2038–48. https://doi.org/10.1016/j.patcog.2006.12.019

[29] Altmann A, Toloşi L, Sander O, Lengauer T. Permutation importance: a corrected feature importance measure. Bioinformatics 2010;26(10):1340–7. https://doi.org/10.1093/bioinformatics/btq134

[30] Rainio O, Teuho J, Klén R. Evaluation metrics and statistical tests for machine learning. Scientific Reports 2024;14(6086):1–12. https://doi.org/10.1038/s41598-024-56706-x

[31] Hossin M, Sulaiman MN. A review on evaluation metrics for data classification evaluations. Int J Data Min Knowl Manag Process Mar 2015;5(2):1–11. https://doi.org/10.5121/ijdkp.2015.5201

[32] Vujović Ž. Classification model evaluation metrics. Int J Adv Comput Sci Appl Jul 2021;12(6):1–8. https://doi.org/10.14569/IJACSA.2021.0120670

[33] Sagir N, Tugcu ZH. Machine-learning-based path loss prediction for vehicle-to-vehicle communication in highway environments. Appl Sci 2024;14(17):7545. https://doi.org/10.3390/app14177545

[34] Wireless C. MK5 OBU. Cohda Wireless Web sitesi. Available online:. https://www.cohdawireless.com/solutions/hardware/mk5-obu/ [accessed on 27 July 2024].

[35] Wu D, Zhu G, Ai B. Application of artificial neural networks for path loss prediction in railway environments. In: Proc. IEEE Int. Conf; 2011. p. 1–6.

[36] Morocho-Cayamcela ME, Maier M, Lim W. Breaking wireless propagation environmental uncertainty with deep learning. IEEE Trans Wireless Commun Aug 2020;19(8):5075–85. https://doi.org/10.1109/TWC.2020.2986202

Author biography



Nuğman Sağır is a Research Assistant in the Department of Electronics and Communication Engineering at Karadeniz Technical University, Trabzon, Türkiye. He received his B.Sc. degree in electrical and electronics engineering from Karadeniz Technical University in 2021 and is currently pursuing his M.Sc. degree in the same field at the Institute of Science. His research focuses on communication systems, machine learning, propagation modeling, and V2V communication. He has contributed to several publications, including studies on machine-learning-based path loss prediction in V2V communication environments.



Zeynep Hasirci Tugcu received her B.S. degree in electrical and electronics engineering in 2008, her M.Sc. degree in electronics engineering in 2011, and her Ph.D. degree in 2017, all from Karadeniz Technical University, Trabzon, Türkiye. She is currently an Associate Professor in the Department of Electrical and Electronics Engineering at Karadeniz Technical University. Her work in propagation modeling and V2V communication has led to several journal and conference publications. Her research interests include communication systems, mobile and satellite communication, propagation modeling, and V2V communication.