Journal of Wireless Networks and Communication Systems
Vol. 2, Issue. 3, Jul-Sep 2026, pp. 50~60
ISSN: 3107-6432, DOI: 10.32595/jwncs/v2i3.2026.39

RAXGBoost: A Radio-Aware Gradient Boosting Framework for Accurate Radio Environment Classification in Wireless Networks

Tripti Dewangan[1], Hemlata Dewangan[1]

[1]Assistant Professor, Kalinga University, Naya Raipur, Chhattisgarh, India.

Article Info

Article History:

Received Jul 06, 2026
Revised Aug 10, 2026
Accepted Sep 11, 2026

Keywords:

Radio Environment Classification,
Line-of-Sight (LoS),
Non-Line-of-Sight (NLoS),
Extreme Gradient Boosting (XGBoost)

ABSTRACT

Accurate classification of the radio environment is essential for understanding wireless propagation conditions and for supporting reliable wireless network management. This paper presents RAXGBoost, a radio feature-based categorization framework based on Extreme Gradient Boosting (XGBoost) that differentiates Line-of-Sight (LoS) from Non-Line-of-Sight (NLoS) propagation conditions utilizing four channel characteristics: path loss (PL), delay spread (DS), angular spread of arrival (ASA), and angular spread of departure (ASD). The structure is evaluated based on 3,284 samples from a 4.85 GHz multiple-input multiple-output (MIMO) channel-sounding dataset gathered using three urban measuring routes, using stratified five-fold cross-validation, and compared with standard Logistic Regression and Interaction-Augmented Logistic Regression under similar folds and characteristics. RAXGBoost achieved an accuracy of 95.01%, macro-precision of 93.60%, macro-recall of 95.62%, macro-F1 score of 94.47%, and ROC-AUC of 0.9891, compared with the two logistic regression baselines' accuracy of 90.41% and 90.19%. RAXGBoost outperformed both baselines for all five metrics. These Findings reveal that gradient-boosted decision trees capture nonlinear relationships among radio propagation features that linear and pairwise-interaction models fail to record. Due to the evaluation is predicated on randomly divided samples, further verification on unseen routes and independent measurement campaigns is required to establish generalization beyond the available dataset.

Corresponding Author:

Tripti Dewangan,
Assistant Professor, Kalinga University,
Naya Raipur, Chhattisgarh, India.

1. INTRODUCTION

The rapid progress of fifth-generation (5G) and arising sixth-generation (6G) wireless communication systems has enhanced the demand for a precise understanding of the radio propagation environment [2]. Wireless networks operate in diverse propagation environments in which the transmitted signal is impacted by blockage, reflection, diffraction, scattering and multipath transmission. For channel characterization, channel modeling, and localization, it is

crucial to ascertain if a received signal arrives under Line-of-Sight (LoS) or Non-Line-of-Sight (NLoS) conditions. [3], [4]. The transmitter and the receiver have an unaffected direct path under LoS circumstances. The direct channel is blocked under NLoS conditions, and the received signal is made up of dispersed, diffracted, or reflected components. Thus, precise detection of LoS and NLoS circumstances yields important details about the propagation environment as a whole.

Radio environment classification is closely associated with the statistical Characteristics of wireless channels. Standardized channel models such as 3GPP TR 38.901 explicate key propagation parameters, including path loss, delay spread and the angular spreads at arrival and departure, for different scenarios and frequencies [5]. Path loss (PL) measures the reduction of the transmitted signal across the channel, while delay spread (DS) describes the temporal dispersion of the multipath components. The angular spread of arrival (ASA) and the angular spread of departure (ASD) explain how extensively the multipath components are distributed at different angles at the transmitter and receiver. Previous studies have applied PL, DS, angular parameters, and other propagation descriptors as inputs to machine learning techniques for channel and scenario identification. [6]. Combining these parameters therefore offer a data-driven method of categorizing LoS and NLoS conditions without explicitly reconstructing the surrounding physical environment or depending on image-based environmental information.

Conventional statistical and machine learning approaches, including logistic regression, support vector devices and decision trees, have been widely used to identify LoS and NLoS. Logistic regression is a natural baseline because of its simple structure, efficiency and interpretability. However, it models the class probability as a linear function of the input features, which may not be expressionally adequate to represent nonlinear relationships between propagation parameters. Interaction terms can encode selected feature interactions, but they must be designed manually and cannot be used to indicate arbitrary decision boundaries [6]. Research on LoS/NLoS classification in real-world settings has also demonstrated the usefulness of data-driven techniques for differentiating measured radio signals from propagation [7].

Gradient-boosted decision trees provide a strong alternative [8]. They construct an ensemble one after the other, fixing the errors of the previous ensemble with each new tree. Unlike linear classifiers, tree assemblies can represent nonlinear decision boundaries and characteristic interactions without the user specifying them in advance. The Extreme Gradient Boosting (XGBoost) method introduced by Chen and Guestrin includes regularization and effective optimization to gradient boosting and has become a widely used method for classification and regression on structured data [9]. Its ability to record nonlinear relationships makes it appropriate for radio environment classification from a limited number of channel parameters. However, high accuracy on randomly divided into samples does not assure transferability to new routes, different propagation conditions or independent initiatives for measurement. Random partitioning can overstated performance by placing nearly similar samples in both the training and test sets because to the strong correlation between measurements made along the same route. [10]. As a result, results on randomly divided data must be taken as within-dataset performance.

In order to solve these problems, this paper introduces RAXGBoost, an XGBoost-based framework for radio environment classification using four channel parameters (PL, DS, ASA, and ASD). The Important contributions of this work are as follows:

(1) Compact LoS/NLoS categorization system based on XGBoost and four standard channel parameters, assessed using 3,284 samples from a 4.85 GHz MIMO channel-sounding dataset that was gathered using three different measurement paths.

(2) A methodical comparison with standard Logistic Regression and Interaction-Augmented Logistic Regression using identical input features, folds, and evaluation metrics, which separates the advantage of the nonlinear ensemble learning over linear and pairwise-interaction models.

(3) A clear explanation of the variations in within-dataset performance and generalization to previously unidentified measurement routes, determining the validation required before the deployment of such classifiers.

2. RELATED WORK

Machine learning has been extensively researched for LoS/NLoS identification in ultra-wideband (UWB), vehicular, and MIMO systems. Early research on UWB localization demonstrated that LoS and NLoS circumstances may be successfully distinguished using information taken from the received waveform and enhance ranging accuracy [4]. Since then, UWB channel impulse responses have been directly processed using convolutional neural networks. [11]. Huang et al. [3] demonstrated how machine learning models utilizing channel features can detect LoS and NLoS conditions for MIMO systems in dynamic contexts.

Chakkour et al. [12] examined machine-learning-based LoS/NLoS identification in vehicle-to-vehicle (V2V) communication. The authors generated channel features from the measured received power and the difference between the measured attenuation and the free-space loss, and reviewed Random Forest, support vector machine, and LogitBoost classifiers while considering feature correlation, feature significance, and reduced input sets. Random Forest was most effective, and the accuracy of categorization was enhanced by a balanced training set. However, this research demonstrates that machine learning can identify propagation conditions in a moving-vehicle environment; its setup and features differ from the 4.85 GHz MIMO channel-sounding measurements used in the present work.

Ma et al. [13] suggested a LoS/NLoS recognition method based on the dynamic multi-level optimization of comprehensive features, examining the ways in which various signal properties aid in distinguishing LoS from NLoS conditions. This paper investigates the function of informative features in propagation classification, similar to RAXGBoost. Instead, the current study compares XGBoost with linear and interaction-augmented logistic regression baselines under the same cross-validation folds using four standard channel parameters.

Baldini [14] submitted a sparse data observers (SDO) outlier detection algorithm to mitigate poisoning attacks on machine-learning-based UWB LoS/NLoS classification. The study emphasizes significance of reliable and high-quality data in real-world applications where measurement errors can occur. Although classifier performance can be negatively impacted by altered data, gradient-boosted trees and linear classifiers are not compared.

Qin et al. [15] combined causal backtracking with a residual neural network (ResNet) for LoS/NLoS classification in UWB sensing, utilizing deep feature extraction to identify the reasons for declining readings as well as to categorize propagation circumstances. RAXGBoost, in contrast, uses a tree-based ensemble on structured channel parameters and offers a reduced-weight alternative.

Wu et al. [16] suggested a LoS/NLoS identification-assisted positioning method for 6G indoor MIMO communications, integrating an enhanced weighted method with a propagation-condition identification Random Forest positioning algorithm, and verified it using measured and ray-tracing-based artificial channels. This work demonstrates the value of propagation knowledge

for positioning; RAXGBoost focuses instead on binary radio environment classification from PL, DS, ASA, and ASD.

This study aims to fill two research gaps. First, it is challenging to evaluate the impact of nonlinear modeling when linear and interaction-augmented baselines under identical experimental conditions are not compared with compact classifiers utilizing standard channel parameters. Second, performance assessed through cross-validation at the random sample level does not prove applicability to previously undiscovered paths. Through a controlled comparison of three classifiers, this work fills in the first gap. It also covers the additional validation needed to evaluate the second.

3. PROPOSED RAXGBoost FRAMEWORK

3.1 Dataset Description

The Radio Signal Processing Laboratory at Niigata University contributed the 4.85 GHz MIMO channel-sounding dataset used in the studies [1]. The measurements were performed. outdoors along three urban routes in Yokohama, Japan: routes R3 and R4 follow the urban macro-cell (UMa) scenario, and route R6 follows the urban micro-cell (UMi) scenario as defined in 3GPP TR 38.901 [5]. The double-directional impulse response was measured for each receiver position, and the data-processing scripts included with the dataset were used to calculate the large-scale parameters (path loss, delay spread, and angle spreads) from the estimated multipath components. Only the valid receiver positions released by the dataset providers are used, giving 3,284 samples. Each sample contains four channel parameters, PL (dB), DS (s), ASA (degrees) and ASD (degrees), accompanied by a binary propagation label, LoS_Tag. The labels were not generated by the authors: LoS_Tag is the LoS/NLoS classification of each receiver position included in the dataset, where a position is marked as NLoS otherwise and as LoS if it is part of the route's LoS position list. The dataset contains 1,078 LoS samples (32.8%) and 2,206 NLoS samples (67.2%), so the classes are imbalanced by a ratio of about 1:2. Table 1 lists the number of samples for each route and class, and the four characteristics for each class are compiled in Table 2.

Table 1. Number of samples per measurement route and propagation class

Route	Scenario	LoS	NLoS	Total
R3	UMa	69	846	915
R4	UMa	182	452	634
R6	UMi	827	908	1,735
Total		**1,078**	**2,206**	**3,284**

Table 2. Median (mean) value of each channel parameter for the LoS and NLoS classes.

Feature	LoS	NLoS
PL (dB)	86.39 (84.30)	115.63 (112.77)
DS (μs)	0.172 (0.179)	0.160 (0.239)
ASA (degrees)	66.37 (69.56)	61.23 (64.80)
ASD (degrees)	66.54 (64.38)	32.33 (43.45)

Before training, the selected characteristics and target labels were examined for missing values, invalid entries, inconsistent and redundant records, and data formats. No samples were eliminated since there were no duplicates or missing values. The GPS coordinates (RxIdx, RxLat, and RxLon) and the receiver index were not included in the input features, and the target label was excluded from the input feature matrix to avoid data leakage.

3.2 Problem Formulation

The dataset is represented as

$$D = \{(\mathbf{x}_i, y_i)\}_{i=1}^{N} \tag{1}$$

where $N = 3284$. The input feature vector of the i-th sample is

$$\mathbf{x}_i = [PL_i,\ DS_i,\ ASA_i,\ ASD_i]^T \tag{2}$$

and the propagation label is

$$y_i = \begin{cases} 1, & \text{LoS condition} \\ 0, & \text{NLoS condition} \end{cases} \tag{3}$$

The objective is to develop a predictive classifier y_i from $\mathbf{x_i}$. The four features describe signal attenuation, temporal multipath dispersion and angular dispersion at the receiver and the transmitter.

3.3 Framework Overview

The RAXGBoost framework uses XGBoost to classify radio transmission conditions as LoS or NLoS. Fig. 1 illustrates its workflow. First, the input dataset is prepared and examined. An XGBoost classifier with 300 decision trees receives the four channel parameters; then the validation procedure outlined in Section 4 is used to compare the model with two baseline classifiers. The standard XGBoost method is used as the underlying learner in this full radio-feature classification pipeline, which is called RAXGBoost [9], and it is not claimed that the boosting technique has changed.

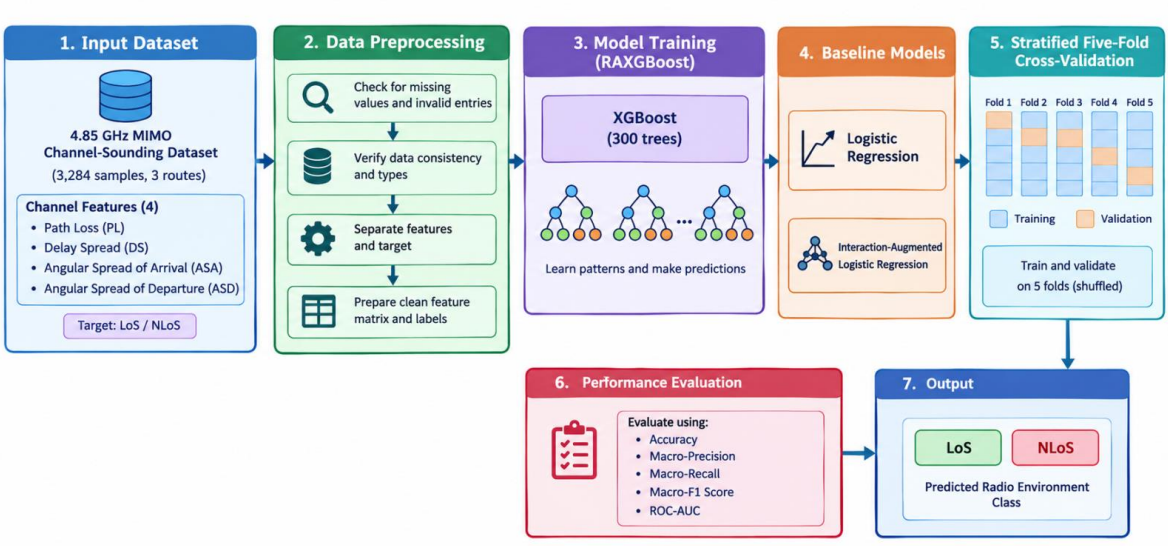


Figure 1. Architecture of the proposed RAXGBoost framework for LoS/NLoS radio environment classification.

3.4 XGBoost-Based Classification

To ascertain the relationship between the propagation class and the channel parameters, XGBoost integrates several decision trees into a nonlinear classification function. The ensemble score is

$$F_K(\mathbf{x}_i) = \sum_{k=1}^{K} f_k(\mathbf{x}_i), f_k \in \mathcal{F} \tag{4}$$

where $K = 300$ is the number of trees, f_k denotes the k-th tree and $\mathcal{F}$ is the space of regression trees. The sigmoid function transfers the ensemble score into the probability of the LoS class:

$$p_i = P(y_i = 1 \mid \mathbf{x}_i) = \frac{1}{1+\exp[-F_K(\mathbf{x}_i)]} \tag{5}$$

and the expected class is acquired with a threshold of 0.5:

$$\hat{y}_i = \begin{cases} 1, & p_i \geq 0.5 \\ 0, & p_i < 0.5 \end{cases} \tag{6}$$

XGBoost minimizes a regularized objective that integrates the classification loss and the model complexity:

$$\mathcal{L} = \sum_{i=1}^{N} l(y_i, p_i) + \sum_{k=1}^{K} \Omega(f_k) \tag{7}$$

where the binary logistic loss is

$$l(y_i, p_i) = -[y_i \log p_i + (1 - y_i)\log(1 - p_i)] \tag{8}$$

and the tree regularization term is

$$\Omega(f_k) = \gamma T_k + \frac{\lambda}{2}\sum_{j=1}^{T_k} w_{kj}^2 \tag{9}$$

Here T_k is the number of leaves in tree k, w_{kj} is the score of leaf j, γ penalizes additional leaves and λ controls L2 consistency of the leaf scores. During training, every iteration, the ensemble is modified additively by one tree:

$$F_t(\mathbf{x}_i) = F_{t-1}(\mathbf{x}_i) + \eta f_t(\mathbf{x}_i) \tag{10}$$

where F_{t-1} is the existing ensemble, f_t is the tree added at iteration t and η is the learning rate. Training stops when K boosting iterations have been completed. Feature standardization is not necessary for XGBoost as it employs decision trees.

3.5 Hyperparameter Configuration

Table 3 lists the XGBoost hyperparameters utilized in every experiment.

Table 3. XGBoost hyperparameters used in every experiment

Hyperparameter	Value
Number of trees (K)	300
Learning rate (η)	[0.3]
Maximum tree depth	[6]
Minimum child weight	[1]
Subsample ratio	[1.0]
Column subsample ratio	[1.0]
L2 regularization (λ)	[1]
Minimum split loss (γ)	[0]
Objective	binary:logistic
Random seed	42

3.6 Baseline Classifiers

To evaluate the advantage of nonlinear ensemble learning, RAXGBoost is compared with standard Logistic Regression and Interaction-Augmented Logistic Regression. Standard Logistic Regression calculates the probability of the LoS class from a linear combination of the input features:

$$p_i = \frac{1}{1+\exp[-(\beta_0+\boldsymbol{\beta}^T\mathbf{x}_i)]} \tag{11}$$

where β_0 is the intercept and $\boldsymbol{\beta}$ is the vector of model coefficients. Interaction-Augmented Logistic Regression combines the six pairwise products of the four (standardized) input characteristics, resulting in the enhanced feature vector

$$\phi(\mathbf{x}_i) = [PL_i,\ DS_i,\ ASA_i,\ ASD_i,\ PL_iDS_i,\ PL_iASA_i,\ PL_iASD_i,\ DS_iASA_i,\ DS_iASD_i,\ ASA_iASD_i]^T \tag{12}$$

with the appropriate probability

$$p_i = \frac{1}{1+\exp\left[-\left(\beta_0+\boldsymbol{\beta}^T\phi(\mathbf{x}_i)\right)\right]} \tag{13}$$

Statistics calculated just on the training folds are used to normalize the features for both logistic regression models to zero mean and unit variance; for the interaction-augmented model, the ten-dimensional vector that results from computing the pairwise products of the standardized features is then standardized once more. To address the 1:2 class imbalance, both models employ L2 regularization with C=1, the L-BFGS solver, and class weights that are inversely proportional to the class frequencies (class_weight = balanced). The baselines were implemented with scikit-learn [17]. The input features, folds, and assessment metrics used by all three classifiers are identical.

4. EXPERIMENTAL PROTOCOL

4.1 Stratified Five-Fold Cross-Validation

Stratified five-fold cross-validation is used to assess within-dataset classification performance. The dataset is divided into five mutually exclusive folds that approximately maintain the initial class distribution:

$$D = \bigcup_{j=1}^{5} D_j, D_j \cap D_m = \varnothing \text{for } j \neq m \tag{14}$$

Every iteration, four folds are used for training and the remaining fold for validation:

$$D_{\text{train}}^{(j)} = D \backslash D_j, D_{\text{val}}^{(j)} = D_j \tag{15}$$

For repeatability, the data are shuffled using a fixed random seed (42), and each sample is forecasted exactly once by a model that was not trained on it (out-of-fold prediction). The same folds are used for all classifiers, and the evaluation metrics are calculated using the pooled out-of-fold predictions.

4.2 Performance Evaluation Metrics

Performance is determined using accuracy, macro-precision, macro-recall, macro-F1 score, and the area under the receiver operating characteristic curve (ROC-AUC). TP, TN, FP, and FN stand for true positives, true negatives, false positives, and false negatives when LoS is the positive class. Accuracy is

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{16}$$

For each class $c \in \{\text{LoS,NLoS}\}$, precision and recall are computed by treating c as the positive class:

$$P_c = \frac{TP_c}{TP_c+FP_c}, R_c = \frac{TP_c}{TP_c+FN_c} \tag{17}$$

and the macro-averaged metrics are the unweighted means over the two classes:

$$\text{Macro-}P = \frac{1}{2}\sum_c P_c \text{ ,Macro-}R = \frac{1}{2}\sum_c R_c \text{ ,Macro-}F1 = \frac{1}{2}\sum_c \frac{2P_cR_c}{P_c+R_c} \tag{18}$$

When the classes are unbalanced, macro-averaging is vital because it provides the LoS and NLoS classes equal weight. The ROC curve plots the true positive rate (TPR) against the false positive rate (FPR) over all classification thresholds, and

$$\text{AUC} = \int_0^1 \text{TPR}\,(\text{FPR})\, d(\text{FPR}) \tag{19}$$

5. RESULTS AND DISCUSSION

5.1 Stratified Cross-Validation Results

Using the identical input features and folds for each model, Table 4 and Fig. 2 compare RAXGBoost with the two logistic regression baselines under stratified five-fold cross-validation.

Table 4. Performance comparison of RAXGBoost and baseline classifiers below stratified five-fold cross-validation.

Model	Accuracy (%)	Macro-Precision (%)	Macro-Recall (%)	Macro-F1 (%)	ROC-AUC
Logistic Regression	90.41	88.50	91.20	89.53	0.9706
Interaction-Augmented Logistic Regression	90.19	88.27	91.18	89.33	0.9728
RAXGBoost (Proposed)	95.01	93.60	95.62	94.47	0.9891

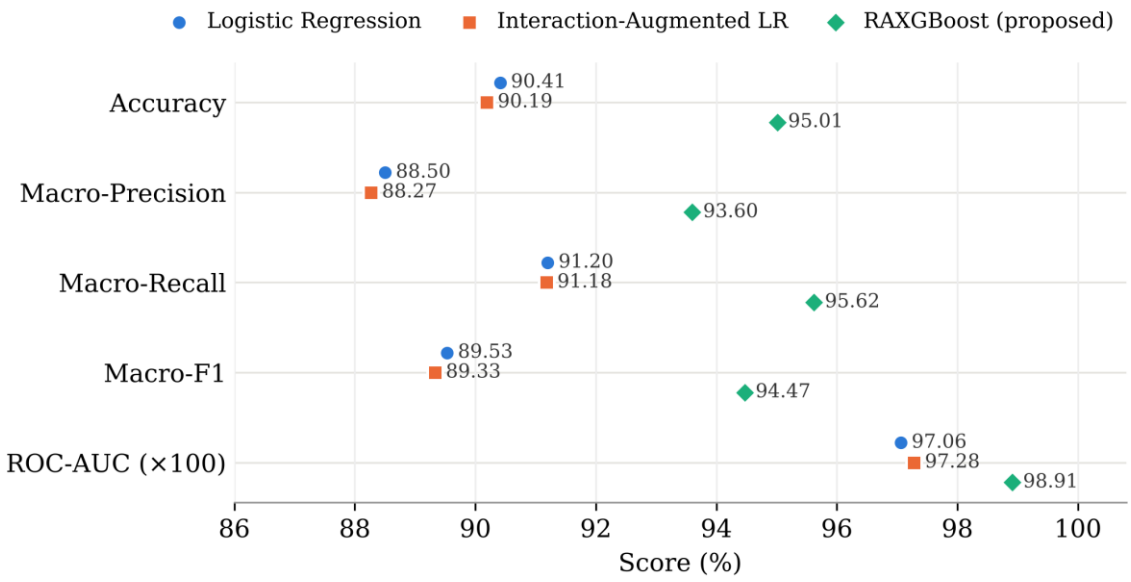


Figure 2. Using stratified five-fold cross-validation, RAXGBoost and the logistic regression baselines are compared across all five assessment criteria (ROC-AUC displayed ×100).

RAXGBoost produced the best results across all metrics. Compared to normal Logistic Regression (90.41%) and Interaction-Augmented Logistic Regression (90.19%), its accuracy of 95.01% is 4.60 and 4.82 percentage points higher, respectively. RAXGBoost raises fewer false alarms and overlooks fewer instances of each class, as evidenced by its macro-precision of 93.60% exceeding the baselines by 5.10 and 5.33 percentage points, and its macro-recall of 95.62% exceeding them by 4.42 and 4.44 percentage points. Its macro-F1 score of 94.47%, which is 4.94 and 5.14 percentage points higher, shows that the two propagation classes have a better balance between recall and precision. Better separation of LoS and NLoS samples across all classification thresholds is indicated by its ROC-AUC of 0.9891, which is 0.0185 and 0.0163 higher than the baselines (0.9706 and 0.9728). Logistic regression was not improved by adding pairwise interaction terms; the interaction-augmented model's performance was nearly comparable to that of the standard model. This implies that nonlinear decision limits and higher-order feature interactions—which are not recorded by straightforward pairwise products—are the sources of RAXGBoost's advantage.

5.2 Discussion

The performance gain of RAXGBoost is consistent with the ability of gradient-boosted decision trees to represent nonlinear decision boundaries and interactions among path loss, delay spread, and angular spread. The most distinctive characteristic in this dataset is path loss, with a median path loss of 115.6 dB for NLoS places and 86.4 dB for LoS positions (Table 2). The other characteristics are less obvious. The angular spreads are greater for LoS positions, particularly the ASD (median 66.5° versus 32.3°), and the median delay spread is nearly the same for both classes; however, NLoS positions exhibit a bigger tail of huge delay spreads. These class disparities also differ between routes; for instance, the ASD is much higher on the UMi route R6 than on the UMa routes for both classes, indicating that the majority of LoS samples originate from R6, which contributes to the total ASD difference. Consequently, on all routes, no single linear combination of the attributes distinguishes LoS from NLoSA tree-based model outperforms linear baselines because it can combine path loss thresholds with conditions on the latency and angular spreads. The ROC-AUC of 0.9891 indicates strong class-ranking performance on the pooled out-of-fold

predictions. However, these metrics represent internal cross-validation performance on pooled out-of-fold predictions and do not, in isolation, establish equivalent performance on unseen routes or other deployment environments.

6. CONCLUSION

This paper offered RAXGBoost, an XGBoost-based framework that categorizes Line-of-Sight and Non-Line-of-Sight radio propagation requirements from four channel parameters: path loss, delay spread, angular spread of arrival, and angular spread of departure. Below, stratified five-fold cross-validation on 3,284 samples from a 4.85 GHz MIMO channel-sounding dataset, RAXGBoost achieved an accuracy of 95.01%, macro-precision of 93.60%, macro-recall of 95.62%, macro-F1 score of 94.47%, and ROC-AUC of 0.9891, exceeding the standard and interaction-augmented logistic regression on every metric. These findings demonstrate that radio propagation conditions can be accurately classified from typical channel data using a small, tree-based model. Future research will assess the framework on previously unexplored paths because these results were acquired using random cross-validation and independent initiatives for measurement at other frequencies and environments, examine its real-time application in network management systems, and add more channel descriptors.

REFERENCES

[1] Radio Signal Processing Laboratory, Niigata University. (2026). 4.85 GHz MIMO channel measurement dataset and Sub-6 GHz data processing manual (Ver. 2.0) [Data set]. Niigata University, Japan.

[2] Rappaport, T. S., Sun, S., Mayzus, R., Zhao, H., Azar, Y., Wang, K., Wong, G. N., Schulz, J. K., Samimi, M., & Gutierrez, F. (2013). Millimeter wave mobile communications for 5G cellular: It will work! IEEE Access, 1, 335–349.

[3] Huang, C., Molisch, A. F., He, R., Wang, R., Tang, P., Ai, B., & Zhong, Z. (2020). Machine learning-enabled LOS/NLOS identification for MIMO systems in dynamic environments. IEEE Transactions on Wireless Communications, 19(6), 3643–3657.

[4] Maranò, S., Gifford, W. M., Wymeersch, H., & Win, M. Z. (2010). NLOS identification and mitigation for localization based on UWB experimental data. IEEE Journal on Selected Areas in Communications, 28(7), 1026–1035.

[5] 3GPP. (2022). Study on channel model for frequencies from 0.5 to 100 GHz (3GPP TR 38.901, Version 17.0.0, Release 17). 3rd Generation Partnership Project.

[6] Alejos, A. V. (2026). Low-dimensional wideband channel descriptors for antenna-aware scenario classification at 28 GHz. IEEE Antennas and Wireless Propagation Letters.

[7] Zrafi, R., Nedil, M., Chehri, A., & Gharsallah, A. (2026). A deep learning-based method for LOS/NLOS detection in harsh underground mining. IEEE Open Journal of Antennas and Propagation.

[8] Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. Annals of Statistics, 29(5), 1189–1232.

[9] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785–794).

[10] Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., & Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography, 40(8), 913–929.

[11] Bregar, K., & Mohorčič, M. (2018). Improving indoor localization using convolutional neural networks on computationally restricted devices. IEEE Access, 6, 17429–17441.

[12] Chakkour, Y., Peñarrocha, V. M. R., Rubio, L., Fernández, H., & Reig, J. (2026). Machine learning-based method for LOS/NLOS identification in V2V environment. IEEE Access, 14, 4189–4207.

[13] Ma, Z., Deng, Z., Tian, Z., Zhang, Y., Wang, J., & Guo, J. (2025). A line of sight/non line of sight recognition method based on the dynamic multi-level optimization of comprehensive features. Sensors, 25(2), 304.

[14] Baldini, G. (2025). On the application of a sparse data observers (SDOs) outlier detection algorithm to mitigate poisoning attacks in ultrawideband (UWB) line-of-sight (LOS)/non-line-of-sight (NLOS) classification. Future Internet, 17(2), 60.

[15] Qin, L., Shi, M., Li, J., & Gu, X. (2025). LOS/NLOS classification using causal backtracking and ResNet in UWB sensing. Physical Communication, 72, 102714.

[16] Wu, Y., Wang, C.-X., Huang, C., Li, J., Zhang, L., Aggoune, E. H. M., & Chen, Y. (2025). A novel LoS/NLoS identification-assisted positioning method for 6G indoor MIMO communications. IEEE Transactions on Communications.

[17] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12, 2825–2830.