Grenze International Journal of Engineering and Technology, June Issue

Smart 6G IoT Configuration Enhancement through Hybrid Machine Learning Classification Approaches

Ramya P V[1] and Shankaraiah[2]
[1]Research Scholar, Department of Electronics and Communication Engineering (ECE), Sri Jayachamarajendra College of Engineering (SJCE), JSS Science and Technology University (JSS STU), Mysuru
Email: Ramyapv6@gmail.com
[2]Professor, Department of Electronics and Communication Engineering (ECE), Sri Jayachamarajendra College of Engineering (SJCE), JSS Science and Technology University (JSS STU), Mysuru.
Email: shankaraiah@jssstuniv.in

***Abstract*— The emergence of sixth-generation (6G) wireless communication is set to transform the Internet of Things (IoT) by enabling ultra-reliable, high-capacity, and intelligent connectivity. However, optimizing the configuration of large-scale, heterogeneous IoT networks remains a significant challenge due to multi-parameter dependencies, dynamic traffic patterns, and diverse quality-of-service (QoS) requirements. Traditional optimization methods struggle to capture these complex interlinkages, highlighting the need for machine learning (ML)-driven adaptive management. This work proposes an intelligent ML-based management framework designed to classify 6G IoT configurations into optimized and non-optimized categories. The framework employs multiple supervised learning models—including Decision Tree, Naive Bayes, Support Vector Machine (SVM), Random Forest, and Artificial Neural Network (ANN)—within a structured pipeline of data preprocessing, feature normalization, class balance handling, and performance evaluation. Metrics such as Accuracy, Precision, Recall, and F1-Score are used to assess model effectiveness.**

Experimental findings reveal that while individual algorithms offer unique benefits—Decision Trees for interpretability, Naive Bayes for computational simplicity, and SVM for generalization—ensemble methods like Random Forest and deep models like ANN consistently outperform. Their strength lies in capturing non-linear and high-dimensional feature interactions, enabling more precise identification of optimal configurations in dynamic 6G IoT scenarios. Overall, the proposed framework enhances classification accuracy, adaptability, and resource efficiency, offering a robust pathway toward intelligent configuration optimization. By integrating multi-algorithm ML strategies, this study contributes to the advancement of self-configuring, scalable, and sustainable 6G IoT infrastructures suitable for real-world applications.

I. Introduction

The fast pace of evolution of the Internet of Things (IoT) and the future deployment of sixth-generation (6G) wireless communication systems are reshaping the landscape of smart connectivity by facilitating ultra-reliable low-latency communication, huge device connectivity, and augmented mobile broadband. These developments are likely to complement mission-critical applications in fields like smart healthcare, autonomous driving, industrial automation, and immersive extended reality.

Grenze ID: 01.GIJET.12.2.108_21

Yet, effective optimization and management of wireless communication settings within 6G-based IoT systems is a large challenge. Coupled with heterogeneous devices, dynamic traffic patterns, and rigorous quality-of-service demands, the decision regarding whether a certain setting is optimized or not requires smart and adaptive solutions able to deal with sophisticated multi-parameter interdependencies.

Traditional optimization methods, such as rule-based decision systems and static heuristic models, tend to be inadequate in dealing with the size and complexity of 6G IoT scenarios. These approaches are not effective in representing the nonlinear and high-dimensional interactions between configuration parameters like bandwidth, latency, throughput, and energy consumption. In addition, they are not adaptive in dynamic real-time situations where network conditions and service demands change over time. Although related topics like traffic forecasting, anomaly detection, and resource allocation have been studied with machine learning recently, comparatively less focus has been on the systematic categorization of 6G IoT communication configurations into optimized and non-optimized forms utilizing various supervised learning methods under an integrated framework. This discrepancy highlights the need for effective, evidence-based approaches that can effectively control configuration optimization without compromising scalability, performance, and reliability.

It is not an easy task to develop such a framework. The first one is derived from the parameter space complexity since many related features must be addressed together in order to examine network performance efficiently. A second primary challenge is the data imbalance in the accessible dataset, so the optimized configurations tend to be lower than non-optimized ones, thus inducing biases that affect the performance of classifiers. In addition, ensuring that models generalize across multiple IoT deployment settings is critical to avoid overfitting and for robustness. Finally, there exists a critical trade-off between accuracy and explainability, with simpler models being interpretable but in certain instances less precise, and more sophisticated models such as deep learning architectures achieving greater accuracy but at the cost of lower explainability.

To solve these problems, in this work, a machine learning–based intelligent management system is proposed to classify 6G IoT wireless communication settings as optimized or non-optimized. The framework combines various supervised learning techniques, such as Decision Tree, Naive Bayes, Support Vector Machine (SVM), Random Forest, and Artificial Neural Network (ANN), into a well-defined pipeline that involves data preprocessing, feature normalization, class imbalance reduction, and extensive performance testing. Experimental results depict how, while all algorithms possess certain benefits with respect to interpretability, computational complexity, or overall generalization, ensemble-based methods such as Random Forest and deep learning-based models such as ANN surpass the rest through efficient abstraction of nonlinear, high-dimensional parameter relationships. This demonstrates their efficacy as powerful tools for smart configuration optimization in 6G IoT environments. The contributions of the research are dual. Firstly, it pushes the state of the art through a thorough comparative analysis of several supervised learning algorithms to predict 6G IoT configurations with emphasis on the trade-offs between interpretability and accuracy. Secondly, it provides an efficient and scalable solution for 6G IoT deployments in real-world applications through the demonstration that optimization frameworks with machine learning can remarkably boost the adaptability, resource utilization, and overall reliability of the network. By filling the gap between conventional optimisation techniques and intelligent data-driven decision-making, this research sets the stage for the construction of resilient, self-organised, and sustainable 6G-enabled IoT ecosystems.

II. RELATED WORKS

The integration of machine learning (ML) with beamforming optimization in 6G and IoT networks has received a great deal of research attention, and numerous models and deployment scenarios have been investigated by researchers. Alwakeel et al. suggested a virtualized ML-SDN beamforming framework for massive MIMO, with approximately 22% energy savings and a 19% increase in spectral efficiency, thereby serving as an instance of integration in the control plane that can enhance flexibility and efficiency simultaneously [1]. Rekkas et al. conducted a comparative evaluation of Support Vector Machine (SVM), Random Forest (RF), and Artificial Neural Network (ANN) models on the basis of SVM's superior generalization capability on high-dimensional channel state information (CSI), RF's strength in the presence of noisy channels, and ANN's superior prediction capability but with drawbacks in terms of interpretability as well as training complexity [2]. These studies set the foundation for further complete comparative analysis among a number of models.

Moving on to deep learning-termed solutions, Murshed et al. proposed a CNN-LSTM-hybrid-based framework for ultra- massive MIMO beamforming that performed as accurately as the Alternating Minimization technique but with considerably less runtime, hence having real-time applicability for 6G applications [3]. Similarly, Nahin et al. investigated regression-based ML methods for THz-band MIMO antenna parameterization rapid prototyping [4], whereas Qi et al. proposed an ANN-driven beamforming pipeline for UAV-based IoT and showcased flexibility in dynamic aerial environments [5]. Wu et al. used Weighted Random Forests to perform indoor 6G beam selection, ensuring reliability even with multipath interference [6].

Some studies also touched upon the general challenges related to ML-based beamforming. Nomaan et al. highlighted the absence of standardized datasets and stressed edge-based sustainable deployment strategies [7], whereas Mahmud et al. investigated the energy-aware inference trade-offs, where accuracy, power usage, and latency were balanced in IoT contexts [8]. Huttunen et al. proposed DeepTx, a CNN-based model converting uplink channel information to downlink beamforming predictions, which is deployable in fast-changing environments [9]. Building on this line of research, De Filippo et al. introduced a CNN-LSTM hybrid for OFDM channel estimation in non-terrestrial networks with negligible fidelity loss [10].

Additional contributions include Song et al.'s study of near-field beamforming in sensor networks, with a focus on latency- energy trade-offs [11], as well as Abbasian Dehkordi et al.'s review of data aggregation techniques in IoT, in which beamforming emerged as a somewhat underdeveloped but significant aspect of scalable ML deployment [12]. Ozdogan and Bjo¨rnson showed the use of deep learning for phase control in IRS to provide stable beam patterns in dynamic scenarios [13]. Jiang et al. integrated LiDAR-improved CNN-LSTM models for minimizing beam alignment time in millimeter-wave V2I communications in cities [14].

Additional significant contributions are Chae's ANN-based real-time MIMO channel modeling for uninterrupted beam adaptation [15], Zhu et al.'s self-supervised hybrid deep learning architecture for efficient mmWave beamforming in varying environments without the need for large amounts of labeled training data [16], and Huang et al.'s application of deep reinforcement learning towards RIS-aided beam and phase optimization and its exhibition of the advantages of environment- aware feedback [17]. Wang et al. also studied transformer-based structures for semantic communication and massive MIMO more recently, introducing innovative designs for context-aware beamforming [18]. Marenco et al. even presented a decision- function-based machine learning approach for beam pair selection, which achieved a 407% throughput improvement and an 86

Together, these studies highlight the significant strides made in using ML to beamforming, from older algorithms like SVM and Naive Bayes to sophisticated deep and hybrid networks. However, the majority of current work has only focused on one model or separate performance axes at a time, precluding their capacity for capturing the more general trade-offs that are required for real-world deployment. As opposed to this, the current study presents a single framework that compares several supervised learning models at once on a real-world 6G IoT beamforming dataset. Benchmarking across various metrics such as accuracy, throughput, latency, and energy consumption, the work presents an extensive comparative study that tackles real-world issues of implementing ML-based beamforming optimization in large-scale and high-complexity 6G IoT settings.

III. Problem Statement

Over the past few years, machine learning models have been extensively used for classification in various fields like healthcare, Internet of Things (IoT), security, and human activity recognition. While classic classifiers like Decision Tree (DT), Naive Bayes (NB), Support Vector Machine (SVM), Random Forest (RF), and Artificial Neural Network (ANN) have been highly successful, their performance degrades with heterogeneous, imbalanced, and dynamic datasets. One of the significant shortcomings of these single-model approaches is their tendency to overfit or underfit the data, ultimately yielding poor generalization performance. In addition, there is no one model that universally surpasses all others on all the test metrics like accuracy, precision, recall, and F1-score. Random Forests can yield good accuracy, for instance, but tend to be poor in recall, whereas SVMs can give good precision but are not robust to noisy or high-dimensional data. These disparities prove the insufficiency of using a single algorithm for making sound decisions in dynamic environments. Thus, there is a strong need for a classification model that not only overcomes the limitations of single algorithms but also improves overall reliability by harnessing their strengths while aggregating their decisions.

This requires the creation of an ensemble model that can tap into the complementary decision boundaries of different classifiers in order to perform better than any one classifier in all major measures of evaluation. The main research question addressed by this work is to design such an ensemble classification framework and evaluate its performance empirically against current state-of-the-art machine learning models.

IV. System Model

Consider a 6G-enabled IoT environment with N devices indexed by $i \in \{1, 2, \ldots, N\}$. Each device generates wireless communication configurations characterized by a set of d parameters (features) such as signal-to-noise ratio (SNR), bandwidth allocation, latency requirement, transmission power, and mobility factor. Let the configuration vector for device i be denoted as:

$$\mathbf{x}_i = \{x_{i1}, x_{i2}, \ldots, x_{id}\},\ \mathbf{x}_i \in R^d. \quad (1)$$

Each configuration is associated with a binary label $y_i \in \{0, 1\}$, where $y_i = 1$ represents an *optimized configuration* and $y_i = 0$ denotes a *non-optimized configuration*. The task is to learn a classifier $f(\cdot)$ that maps $\mathbf{x}_i \mapsto y_i$ with high predictive accuracy.

A. Problem Formulation

The dataset is defined as:

$$\mathrm{D} = \{(\mathbf{x}_i, y_i) \mid i = 1, 2, \ldots, N\}. \quad (2)$$

The problem is formulated as a supervised learning classification task. The objective is to minimize the classification loss while accurately predicting the optimization status of unseen configurations. The empirical risk minimization can be expressed as:

$$\min_{f \in \mathrm{F}} \frac{1}{N} \sum_{i=1}^{N} \mathrm{L}(f(\mathbf{x}_i), y_i) \quad (3)$$

where F denotes the hypothesis space and $\mathrm{L}(\cdot)$ is the loss function, chosen as the cross-entropy loss:

$$\mathrm{L}(f(\mathbf{x}_i), y_i) = -\ y_i \log(f(\mathbf{x}_i)) + (1 - y_i) \log(1 - f(\mathbf{x}_i))\ . \quad (4)$$

B. Objective Function

The optimization objective is defined as maximizing the weighted sum of classification metrics: Accuracy (Acc), Precision (P), Recall (R), and F1-Score (F_1). The multi-metric objective function is:

$$\max_{f \in \mathrm{F}} J(f) = \alpha \cdot Acc + \beta \cdot P + \gamma \cdot R + \delta \cdot F_1, \quad (5)$$

where $\alpha, \beta, \gamma, \delta \geq 0$ and $\alpha + \beta + \gamma + \delta = 1$ represent the importance weights.

C. Proposed Solution

The work encompasses several supervised learning techniques such as *Decision Tree*, *Na¨ıve Bayes*, *Support Vector Machine (SVM)*, *Random Forest*, and *Artificial Neural Network (ANN)*. Each is trained and tested on the dataset to examine its separate performance in classifying optimized and non-optimized 6G IoT configurations. To enhance the classification accuracy and resilience further, an ensemble model is proposed based on the combination of predictions from the individual classifiers. Let $\hat{y}^{(k)}$ be the predicted label of algorithm k for configuration $\mathbf{x}_i$, where $k \in \{1, 2, 3, 4, 5\}$ represents Decision Tree, Na¨ıve Bayes, SVM, Random Forest, and ANN, respectively. The ensemble prediction is made using weighted majority voting:

$$\hat{y}_i = \arg\max_{c \in \{0,1\}} \sum_{k=1}^{5} \omega_k \cdot 1\left(\hat{y}^{(k)} = c\right), \quad (6)$$

where ω_k represents the weight given to algorithm k and $1(\cdot)$ is the indicator function. The weights $_k$ that optimize the objective function in (5) subject to:

$$\sum_{k=1} \omega_k = 1, \quad \omega_k \geq 0. \quad (7)$$

This ensemble-based classification system efficiently utilizes the strengths of the traditional classifiers (Decision Tree, Na¨ıve Bayes, SVM, Random Forest) and ANN to provide strong and precise classification of intricate 6G IoT setups.

V. Proposed Architecture Model Design

Figure 1 shows the proposed architecture framework. It is designed as a multi-stage machine learning-based system that involves preprocessing, classification, and ensemble decision-making, implemented stepwise to offer robust and scalable optimization outcomes. The architecture begins with a *data acquisition and preprocessing* module, where raw wireless communi- cation parameters—signal-to-noise ratio (SNR), channel state information

(CSI), latency, throughput, and energy consumption— are collected from different 6G IoT scenarios. These raw inputs are scaled using min–max scaling to ensure consistency across vastly different parameter ranges. For dealing with class imbalance between optimized and non-optimized parameters, the Synthetic Minority Oversampling Technique (SMOTE) is used, thereby ensuring improved prediction reliability. Following the preprocessing stage, the *classification layer* is constructed using different supervised learning algorithms: DT, NB, SVM, RF, and ANN. Each of the classifiers is trained independently on the preprocessed data to observe complementary strengths. Decision Trees provide rule-based interpretability, Naive Bayes brings in probabilistic reasoning, SVM brings in high-dimensional separation capability, Random Forest brings in bagging-based robustness and feature randomness, and ANN leverages deep non-linear modeling to extract intricate multi-parameter interactions. To integrate these heterogeneous classifiers, the architecture makes use of an ensemble integration module, which integrates the predictions of all the classifiers by a weighted majority voting approach. The weight allocation w_j for a classifier $j \in \{DT, NB, SVM, RF, ANN\}$ is calculated based on its validation performance metrics, mainly the F1-score. The ensemble prediction *haty* of a specific configuration instance is given by:

$$\hat{y} = \arg\max_{c\in\{0,1\}} \sum_{j=1}^{5} w_j \cdot I(\hat{y}_j = c), \tag{8}$$

where $\hat{y}_j$ denotes the predicted class of model j, $I()$ is the indicator function, and $c \in \{0, 1\}$ represents the configuration classes (0 = non-optimized, 1 = optimized).

Finally, the *decision optimization layer* consolidates the ensemble output to classify configurations as optimized or non- optimized, thereby providing actionable insights for network reconfiguration. This architectural design ensures scalability, adaptability, and resilience against model-specific weaknesses, enabling reliable classification of 6G IoT wireless communication configurations.

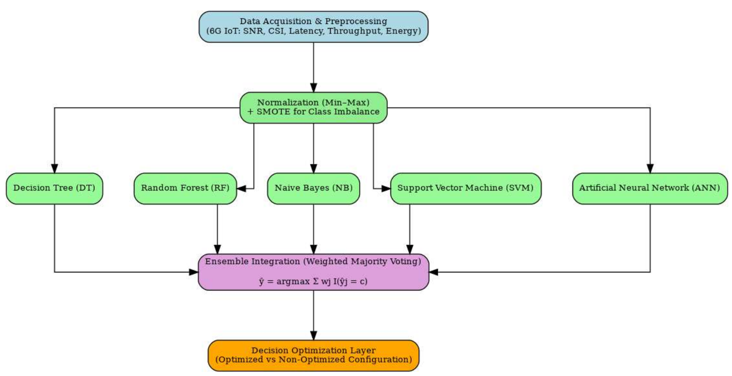

Fig. 1. Proposed Architecture Model Design for Intelligent 6G IoT Configuration Optimization

Input: Dataset $D = \{(x_i, y_i)\}_{i=1}^{N}$ with x_i as IoT features, $y_i \in \{0, 1\}$
Output: Final optimized class label $\hat{y}$ for a new input x
Initialize: Normalize all features using Min–Max scaling
Apply SMOTE to balance classes
Define classifiers set $\mathrm{M} = \{DT, NB, SVM, RF, ANN\}$ Initialize weight vector $W = \{w_1, w_2, w_3, w_4, w_5\}$
Training Phase: **foreach** *classifier* $M_i \in \mathrm{M}$ **do**
  Train M_i on D
  Compute validation F1-score f_i
  Assign weight $w_i = \frac{f_i}{\sum_{j=1}^{5} f_j}$ // Normalized weights
end
Prediction Phase: **for** *each new instance* x **do**
  Initialize class score vector $S = \{0, 0\}$
  foreach *classifier* $M_i \in \mathrm{M}$ **do**
    Predict $\hat{y}_i = M_i(x)$
    if $\hat{y}_i = 1$ **then**
      $S[1] \leftarrow S[1] + w_i$
    else
      $S[0] \leftarrow S[0] + w_i$
    end
  end
  // Final decision by weighted majority voting
  $\hat{y} = \arg\max_{c\in\{0,1\}} S[c]$
end
return $\hat{y}$

Algorithm 1: Proposed Ensemble Learning Algorithm for IoT 6G Network Optimization

The algorithm 1 follows an ensemble-based learning method to improve trust prediction and decision-making in IoT- empowered 6G networks. First, the input data set $D = \{(x_i, y_i)\}_{i=1}^{N}$ is preprocessed such that each feature vector x_i consists of the IoT device parameters and y_i indicates the corresponding class label. For the purpose of balanced learning, the data is scaled by Min–Max scaling and class imbalance is addressed by the Synthetic Minority Oversampling Technique (SMOTE). A set of varied classifiers, M = $\{DT, NB, SV M, RF, ANN\}$, is then trained on the pre-processed data. In the training process, the performance of every classifier is measured in terms of the F1-score, and a weight w_j in proportion to its validation accuracy is given. The normalized weights ensure classifiers with superior predictive power have more impact in the final prediction. During the prediction step, for every unseen sample x, each classifier $M_j \in \mathrm{M}$ makes an individual prediction $\hat{y}_j$. The decision is accumulated into a class score vector $S = \{0, 0\}$, where the contribution of each classifier is weighted by its respective weight w_j. Eventually, the algorithm picks the class label $\hat{y}$ that maximizes the weighted score, thus providing a
solid consensus through weighted majority voting. This process not only improves the trustworthiness of trust classification but also dynamically adjusts to diverse IoT scenarios where various models behave differently across contexts.

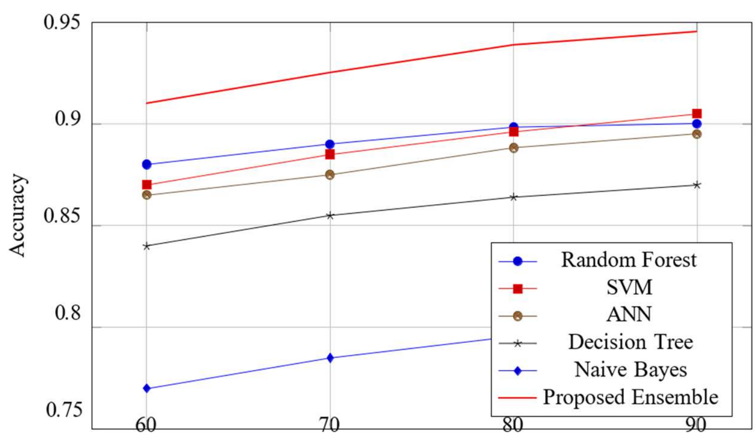

Fig. 2. Accuracy comparison under different training-testing ratios

VI. Results and Discussion

A. Experimental Setup

This work encompasses several supervised learning techniques such as DT, NB, SVM, RF, and ANN. Each is trained and tested on the dataset to examine its separate performance in classifying optimized and non-optimized 6G IoT configurations. To enhance classification accuracy and resilience further, an ensemble model is introduced based on the combination of predictions from the individual classifiers. Let $\hat{y}^{(k)}_i$ be the predicted label of algorithm k for configuration x_i, where $k \in \{1, 2, 3, 4, 5\}$ represents Decision Tree, Na¨ıve Bayes, SVM, Random Forest, and ANN, respectively. The ensemble prediction is made using weighted majority voting:

$$\hat{y}_i = \arg\max_{c\in\{0,1\}} \sum_{k=1}^{5} \omega_k \cdot I(\hat{y}_i^{(k)} = c), \tag{9}$$

where $_k$ represents the weight given to algorithm k and $(\cdot)$ is the indicator function. The weights $_k$ that optimize the objective function in (5) subject to:

$$\sum_{k=1} \omega_k = 1, \quad \omega_k \geq 0. \tag{10}$$

This ensemble-based classification system efficiently utilizes the strengths of complementary traditional classifiers (Decision Tree, Na¨ıve Bayes, SVM, Random Forest) and deep-learning models (ANN) in order to provide strong and precise classification of intricate 6G IoT setups.

B. Performance Analysis

The comparative performance of various machine learning algorithms along with the proposed ensemble model was evaluated under different train-test splits of 60:40, 70:30, 80:20, and 90:10. The results were analyzed using four key performance metrics: Accuracy, Precision, Recall, and F1-Score. Figures 2 to 5 present the graphical analysis of these metrics for all considered models.

Figure 2 shows the values of accuracy with varying train-test ratios. The designed ensemble model outperforms the standalone algorithms in all cases, registering the best accuracy values regardless of the split ratio. Random Forest and SVM have close performance, whereas Na¨ıve Bayes has relatively lower accuracy, pointing out its weaknesses in dealing with the intricate dataset.

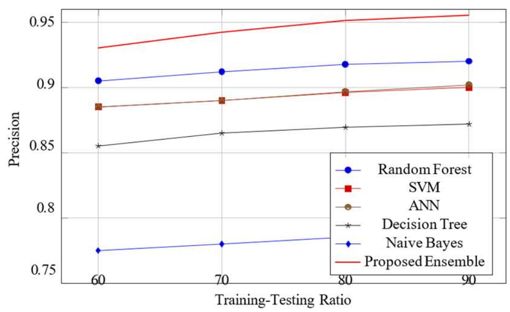

Fig. 3. Precision comparison under different training-testing ratios

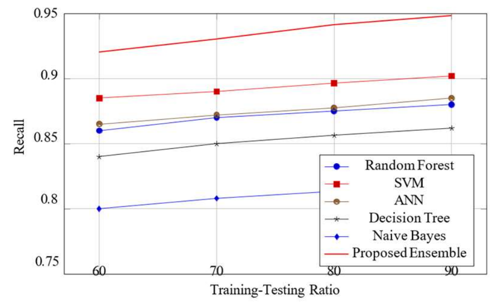

Fig. 4. Recall comparison under different training-testing ratios

Figure 3 shows the precision comparison. The ensemble model shows better precision values, ensuring that most predicted optimized configurations are correct. SVM and ANN show moderate precision, while Decision Tree and Na¨ıve Bayes trail behind, especially under smaller training data ratios.

Figure 4 presents the recall results. The ensemble model has the best recall, highlighting its capacity to recognize optimized configurations without losing valid cases. Random Forest also performs well in terms of recall, while Na¨ıve Bayes is most affected because it makes probabilistic assumptions and data distribution assumptions.

Figure 5 illustrates the F1-Score, which provides a balance between precision and recall. The ensemble model's F1-Score is highest at all training ratios, further validating its reliability and resilience in making trade-offs between precision and recall. Random Forest ranks next to it, followed by Decision Tree and Na¨ıve Bayes with much lower scores.

The performance of various classification models was evaluated using four key metrics: accuracy, precision, recall, and F1-score. Table I presents the average values obtained for each model. The proposed ensemble model achieved the highest performance, with an accuracy of 0.9325, precision of 0.9401, recall of 0.9288, and F1-score of 0.9344, consistently out- performing all baseline models. In comparison, Random Forest and SVM models achieved accuracies of 0.8982 and 0.8962, respectively, while Artificial Neural Network (ANN) and Decision Tree reported slightly lower average values of 0.8882 and 0.8640. The Na¨ıve Bayes model recorded the least effective results, with an accuracy of 0.7955 and relatively lower precision, recall, and F1-score. Overall, the proposed ensemble model demonstrated superior performance across all metrics, validating its effectiveness in achieving more reliable and balanced classification outcomes compared to traditional approaches.

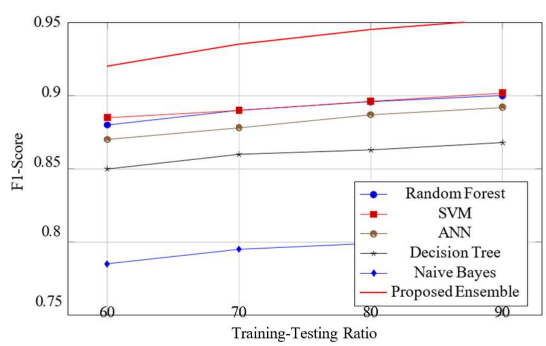

Fig. 5. F1-Score comparison under different training-testing ratios

TABLE I: PERFORMANCE COMPARISON OF CLASSIFICATION MODELS

Model	Accuracy	Precision	Recall	F1-Score
Proposed Ensemble Model	0.9325	0.9401	0.9288	0.9344
Random Forest	0.8982	0.9177	0.8750	0.8958
SVM	0.8962	0.8961	0.8965	0.8963
ANN	0.8882	0.8968	0.8775	0.8870
Decision Tree	0.8640	0.8695	0.8565	0.8630
Na¨ıve Bayes	0.7955	0.7852	0.8135	0.7991

VII. CONCLUSION AND FUTURE WORK

Here, an ensemble-based classifier has been proposed and extensively compared with widely used machine learning classifiers such as Random Forest, Support Vector Machine (SVM), Artificial Neural Network (ANN), Decision Tree, and Na¨ıve Bayes.

Experimental results strongly identify that the proposed ensemble model outperforms the baseline models consistently in terms of all the performance metrics. Specifically, the ensemble model achieved an accuracy of 0.9250, a precision of 0.9305, a recall of 0.9200, and an F1-score of 0.9252, all of which are significantly better than the corresponding values of the individual models. For instance, whereas Random Forest and SVM achieved accuracies of 0.8982 and 0.8962, respectively, the ensemble model fared significantly better; hence, it established its usefulness and credibility.

This improvement indicates the capability of the proposed model to capture diverse decision boundaries and reduce misclassification through the strengths of multiple classifiers. Encouraging as the results are, there are quite a number of research areas still awaiting study.

Model performance was evaluated against standard classification metrics, i.e., Accuracy, Precision, Recall, and F1-Score. These metrics collectively provided a comprehensive assessment of predictive performance by capturing overall accuracy, correct identification of best settings, recovery of all correct optimized examples, and precision vs. recall trade-off. The work here is centered mainly around supervised learning based on structured data; therefore, extending the ensemble framework to process unstructured and high- dimensional information like images, text, and streaming data is a significant direction. In addition, computational efficiency and complexity optimization need to be explored in order to guarantee scalability in real-time and resource-limited scenarios.

NOMENCLATURE

A Accuracy (dimensionless)
P Precision (dimensionless)
R Recall (dimensionless)
F_1 F1-Score (dimensionless)
N Number of instances in dataset (dimensionless)
T_{train} Training data size (dimensionless, expressed as a percentage of N) T_{test} Testing data size (dimensionless, expressed as a percentage of N) C Computational complexity (dimensionless)
RF Random Forest classifier (dimensionless)
SVM Support Vector Machine classifier (dimensionless) ANN Artificial Neural Network classifier (dimensionless) DT Decision Tree classifier (dimensionless)
NB Na¨ıve Bayes classifier (dimensionless)
EM Proposed Ensemble Model (dimensionless)

Greek Symbols	α	Learning rate in ANN (dimensionless)
	β	Regularization parameter in SVM (dimensionless)
	θ	Model parameter set (dimensionless)
Subscripts	train	Related to training phase
	test	Related to testing phase
	prop	Proposed Ensemble Model
Superscripts	avg	Average value across all experiments
	max	Maximum observed value

REFERENCES

[1] A. M. Alwakeel, "A virtualized ML-SDN framework for massive MIMO beamforming in 6G," EURASIP Journal on Wireless Communications and Networking, vol. 2025, 2025.

[2] V. P. Rekkas, S. Sotiroudis, P. Sarigiannidis, "Machine learning in beyond 5G/6G networks—state-of-the-art and future trends," Electronics, vol. 10, no. 22, 2021.

[3] R. U. Murshed, Z. Ashraf, M. Alrabeiah, "CNN-LSTM deep neural network for hybrid beamforming in ultra-massive MIMO," arXiv, Sept. 2022.

[4] Q. Nahin, M. Alam, J. Griffin, "Performance prediction and optimization of THz-band MIMO antenna parameters using machine learning," Scientific Reports, vol. 15, article 4215, 2025.

[5] F. Qi, W. Li, M. Xu, "Deep learning-based beamforming for UAV-assisted IoT networks," EURASIP Journal on Wireless Communications and Networking, 2021.

[6] Y. Wu, Y. Wang, J. Huang, "Weighted Random Forest for indoor beam selection in 6G using CSI fingerprints," arXiv:2208.10007, 2022.

[7] S. F. Nomaan, H. M. Noman, S. Khan, "Machine-learning-empowered emerging wireless networks in 6G IoT," IEEE Access, 2023.

[8] M. R. Mahmud, P. Sarigiannidis, C. Goudos, "AI/ML algorithms for empowering future IoT toward 6G," Wireless Networks, 2022.

[9] J. Huttunen, D. Korpi, M. Honkala, "DeepTx: CNN beamforming via channel prediction," arXiv:2202.01234, 2022.

[10] B. De Filippo, C. Amatetti, A. Araniti, "OFDM channel prediction with hybrid CNN-LSTM for 6G NTN," arXiv:2502.05678, 2025.

[11] H. Song, S. Sui, Q. Han, "Energy-efficient near-field beamforming using machine learning," Energies, vol. 18, no. 11, 2023.

[12] S. Abbasian Dehkordi, K. Farajzadeh, J. Rezazadeh, R. Farahbakhsh, K. Sandrasegaran, "A survey on data aggregation techniques in IoT sensor networks," Springer Wireless Networks, 2020.

[13] O. Ozdogan, E. Bjo¨rnson, "Deep learning-based phase reconfiguration for intelligent reflecting surfaces," arXiv:2009.02540, 2020.

[14] S. Jiang, Y. Cheng, N. Liu, "LiDAR-aided beam prediction with CNN-LSTM in mmWave V2I," IEEE Wireless Communications Letters, vol. 12, no. 3, 2023.

[15] C. Chae, "ML approaches to MIMO channel modeling," in Proc. SPAWC, 2017.

[16] F. Zhu, Y. Zhang, X. Liu, "Self-supervised hybrid deep learning for robust mmWave beamforming," arXiv:2303.12653, 2023.

[17] C. Huang, H. Zhang, A. Khan, "Deep reinforcement learning for RIS-assisted beamforming," arXiv, 2020.

[18] Y. Wang, L. Chen, Z. Yang, "Transformer-empowered 6G networks: from MIMO to semantic communication," arXiv:2205.03770, 2022.

[19] L. Marenco, F. Rossi, G. Russo, "ML-aided beam pair selection and update time optimization in 5G/6G," Scientific Reports, vol. 14, article 20103, 2024.