

Machine Learning for Terahertz Dual-Band Antennas: A Comparative Study of Modeling Approaches

Sharad Shahi^a, Ritu Roumya Samal^b, Prafulla K. Dash^b, Kajal Parashar^b, S.K.S. Parashar^{b,*}

^a*School of Computer Engineering, Kalinga Institute of Industrial Technology (KIIT) Deemed to be University, Bhubaneswar, 751024, Odisha, India*

^b*Nanosensor Lab, School of Applied Sciences, Kalinga Institute of Industrial Technology (KIIT) Deemed to be University, Bhubaneswar, 751024, Odisha, India*

Abstract

This paper explores machine learning for terahertz antennas through a comparative study of modeling approaches. We present the design of a dual-band antenna operating at 7.713 THz and 9.029 THz, featuring a rectangular slotted patch on a polyimide substrate. The antenna exhibits impressive performance metrics, including a return loss of -48.974 dB, a bandwidth of 848.934 GHz (7.389 to 8.238 THz), a gain of 8.102 dB, and a Voltage Standing Wave Ratio (VSWR) of less than 2. Utilizing neural network models, we predict critical performance parameters such as resonant frequency, return loss, and gain, supported by a diverse dataset generated from a comprehensive parametric sweep in CST Microwave Studio. Our analysis reveals that neural networks and ensemble methods significantly outperform traditional modeling approaches, highlighting the transformative potential of machine learning in optimizing antenna designs for wireless communication systems while significantly enhancing performance optimization.

Keywords: Machine learning, Terahertz (THz) Antenna, Dual-band Antenna, Bandwidth, Neural Network, Ensemble Method

1. Introduction

In the pursuit of advanced wireless communication systems, machine learning (ML) has emerged as a pivotal tool for optimizing antenna design, particularly within the terahertz (THz) frequency range [1, 2]. Spanning approximately 0.1 to 20 THz, the THz band bridges the microwave and infrared regions of the electromagnetic spectrum, offering unique opportunities for high-speed data transmission and innovative applications, including imaging systems, spectroscopy, and non-invasive medical diagnostics [3, 4]. As the demand for higher data rates and enhanced channel capacity intensifies, THz communication technologies are becoming increasingly critical. The THz frequency range not only supports advancements in wireless communications but also enables a variety of applications that require rapid data exchange and precision, further underscoring its relevance in modern technology.

Microstrip patch antennas (MSPAs) have emerged as essential components in communication devices, particularly well-suited for THz applications due to their low profile, cost-effectiveness, and design flexibility [5]. The rising demand for compact, multi-band antennas has led to the recognition of MSPAs as a versatile solution, capable of operating across various frequency bands while maintaining ease of fabrication. However, traditional antenna design methodologies frequently fall short of meeting the stringent bandwidth and performance metrics required for these emerging applications [6]. This limitation highlights the need for innovative design approaches that integrate machine learning techniques to optimize antenna performance and address the evolving demands of the wireless communication landscape.

In the last decade, machine learning has profoundly impacted various domains, including communication systems, driven by advancements in computational capabilities and the growing cost effectiveness of ML technologies. Unlike traditional methods, ML capitalizes on large datasets to significantly improve performance,

*Corresponding author
E-mail addresses: sksparashar@yahoo.com, sksparasharfpy@kiit.ac.in (S.K.S. Parashar)

particularly in the design of custom antennas [7]. Conventional antenna design, which often requires manual tuning of parameters, is both labor intensive and time consuming [2, 5]. Inefficient tuning can result in sub-optimal antenna performance, limiting overall operational efficiency. To address these challenges, researchers are increasingly turning to ML algorithms to accelerate and optimize the design process [2]. Without ML, achieving the required speed and precision in antenna design becomes notably difficult [8]. The advantages of ML driven approaches are substantial, offering reductions in design time, enhancements in computational efficiency, lower operational costs, preservation of design feasibility, and minimization of simulation times [5, 7, 8]. Furthermore, maintaining low error rates while ensuring high productivity is challenging without ML [8]. These techniques streamline the tuning of critical antenna parameters and improve the predictive accuracy of essential performance metrics, such as resonant frequency, return loss, and gain [1, 2]. By leveraging the capabilities of machine learning, the antenna design process becomes significantly more efficient and resource effective, addressing the urgent need for rapid advancements in wireless communication technologies [7].

Incorporating machine learning into antenna design, particularly neural networks (NN), has proven to be highly effective in predicting critical performance parameters [1]. This paper presents a novel approach utilizing a neural network model specifically designed to forecast these performance characteristics for a dual-band microstrip patch antenna operating within the THz spectrum. The integration of machine learning not only streamlines the design process but also facilitates the exploration of complex parameter interactions that traditional methods may overlook [5, 9]. Moreover, ensemble techniques, which synthesize predictions from multiple models, have shown considerable effectiveness in enhancing prediction accuracy [5, 9]. This study evaluates various machine learning models, including ensemble methods such as Hist Gradient Boosting, to assess their performance in optimizing complex antenna designs for THz applications. The potential of ML to generate trained models that can be quickly optimized for varying design objectives is particularly noteworthy [7]. These models can yield insights that lead to more refined designs, significantly reducing the time and resources traditionally required for antenna development [5, 7].

Machine learning techniques offer transformative advantages over traditional design methods, enabling rapid prototyping and thorough exploration of complex parameter interactions [2, 5, 9]. These advancements significantly accelerate innovation in antenna design, positioning ML as a cornerstone in future wireless communication systems. Recent advancements in materials science and fabrication techniques further amplify the prospects for integrating machine learning into antenna design, heralding a new era of innovative applications. These developments not only enhance the capabilities of MSPAs but also pave the way for next generation wireless technologies that can meet the increasing demands of connectivity and data transmission in an increasingly digital world [2]. As we continue to explore the intersections between machine learning and antenna design, we stand poised for significant advancements that could reshape the future of communication technology.

This work's significant offerings and objectives are:

1. Integration of Machine Learning for Antenna Design Optimization: This work emphasize on the use of ML techniques in optimizing the performance characteristics of microstrip patch antennas. By leveraging various models, including Neural Networks, Linear Regression, Ridge Regression, Bayesian Ridge, and Support Vector Machines (SVM) with different kernels, this study seeks to understand how ML can predict critical antenna parameters such as resonant frequency, return loss, and gain. The aim is to offer an innovative approach that can reduce the time and computational effort required for antenna design by traditional methods. This approach represents a significant departure from traditional antenna design methods, which rely heavily on empirical formulas, trial and error, and extensive simulations. By leveraging ML, the design process becomes more efficient and adaptable to various design constraints and requirements.
2. Comparative Analysis of Machine Learning Models: A thorough comparison of different ML models is conducted using the same dataset, scaling, and train-test split. The study examines the performance of these models in predicting antenna characteristics, offering insights into their relative effectiveness. This comparison is vital in identifying the most suitable ML models for specific antenna design tasks.
3. Ensemble Methods for Enhanced Prediction Accuracy: The study also explores the application of ensemble methods, specifically focusing on decision trees, hist gradient boosting and random forest

models. These methods are investigated for their ability to enhance prediction accuracy by combining the strengths of multiple models. The objective here is to determine if ensemble methods can outperform individual models in terms of predictive performance, thereby providing a more robust and reliable tool for antenna design.

4. Contribution to the Field of Antenna Design: The research contributes significantly to the field of antenna design by demonstrating how modern ML techniques can be effectively integrated into the design process. The findings of this study are expected to pave the way for more advanced and efficient antenna design methodologies, potentially leading to innovations in wireless communication systems.
5. Exploration of the Future of Machine Learning and AI in Engineering: The study also reflects on the broader implications of the upcoming era of ML and AI in engineering. By showcasing the potential of these technologies in optimizing complex engineering designs, this work highlights the transformative impact that ML and AI are expected to have in various engineering domains, including antenna design.

The following sections of this paper provide a clear overview of the antenna design and configuration, highlighting the innovative methods used to improve performance. We begin with the results and performance analysis of the antenna design, focusing on key findings. Next, we explore machine learning techniques for optimizing antenna design, evaluating the outcomes of the proposed neural network model. We then compare various machine learning models to assess their effectiveness in antenna optimization. Additionally, we discuss ensemble techniques that enhance performance predictions, drawing important insights from this analysis. Finally, the paper concludes with a summary of the findings and their importance in the field of antenna design.

2. Antenna design and configuration

The antenna's radiating surface and ground plane are constructed from copper, selected for its superior electrical conductivity and cost effectiveness, crucial for achieving optimal signal transmission efficiency. The substrate material is polyimide, characterized by a relative permittivity of 3.5 and a minimal dissipation factor of 0.006, making it suitable for THz applications [10].

Given that the antenna is intended for operation in the THz range (0.1 to 20 THz), all measurements are obtained in the micrometer (μm) range to align with the wavelength of terahertz radiation. The initial design begins with a rectangular patch of dimensions $28 \mu\text{m} \times 22 \mu\text{m} \times 0.4 \mu\text{m}$. Symmetry is a critical element in the design, which incorporates a spear like structure to optimize the radiation pattern. This geometric configuration not only ensures balanced distribution of electromagnetic waves but also improves overall efficiency in signal propagation.

2.1. Design Methodology

The design methodology for this rectangular microstrip patch antenna focuses on achieving optimal symmetry and performance for THz frequency applications through systematic refinement and machine learning driven optimization. The use of copper as both the radiating patch and ground plane material, combined with the intricate design modifications, results in an antenna that is not only efficient but also cost effective. Copper was selected as the material for both the patch and ground plane of the antenna due to its exceptional electrical conductivity, which ensures efficient current flow and enhances radiation characteristics. This choice minimizes resistive losses, contributing to overall antenna performance and making it ideal for high frequency applications. A key feature of the design is the incorporation of a 50Ω inset feeding mechanism, which ensures efficient power transfer and optimal impedance matching. Complementing the choice of material, the substrate material is also critical for optimizing microstrip patch antenna (MSPA) designs, particularly in the terahertz (THz) range [1]. Polyimide was chosen for its favorable dielectric properties, featuring a relative permittivity of 3.5 and a low dissipation factor of 0.006 [10]. These characteristics are crucial for minimizing signal loss and maintaining stable performance at elevated frequencies. The substrate dimensions were set at $50 \mu\text{m} \times 41 \mu\text{m} \times 3 \mu\text{m}$ to effectively manage fringing effects that can significantly influence radiation behavior within the THz spectrum. Compared to other common substrates like FR-4, which tends to exhibit instability at THz frequencies, polyimide provides enhanced stability and reduced signal degradation [3]. This combination of copper for the patch and ground, along with polyimide as the

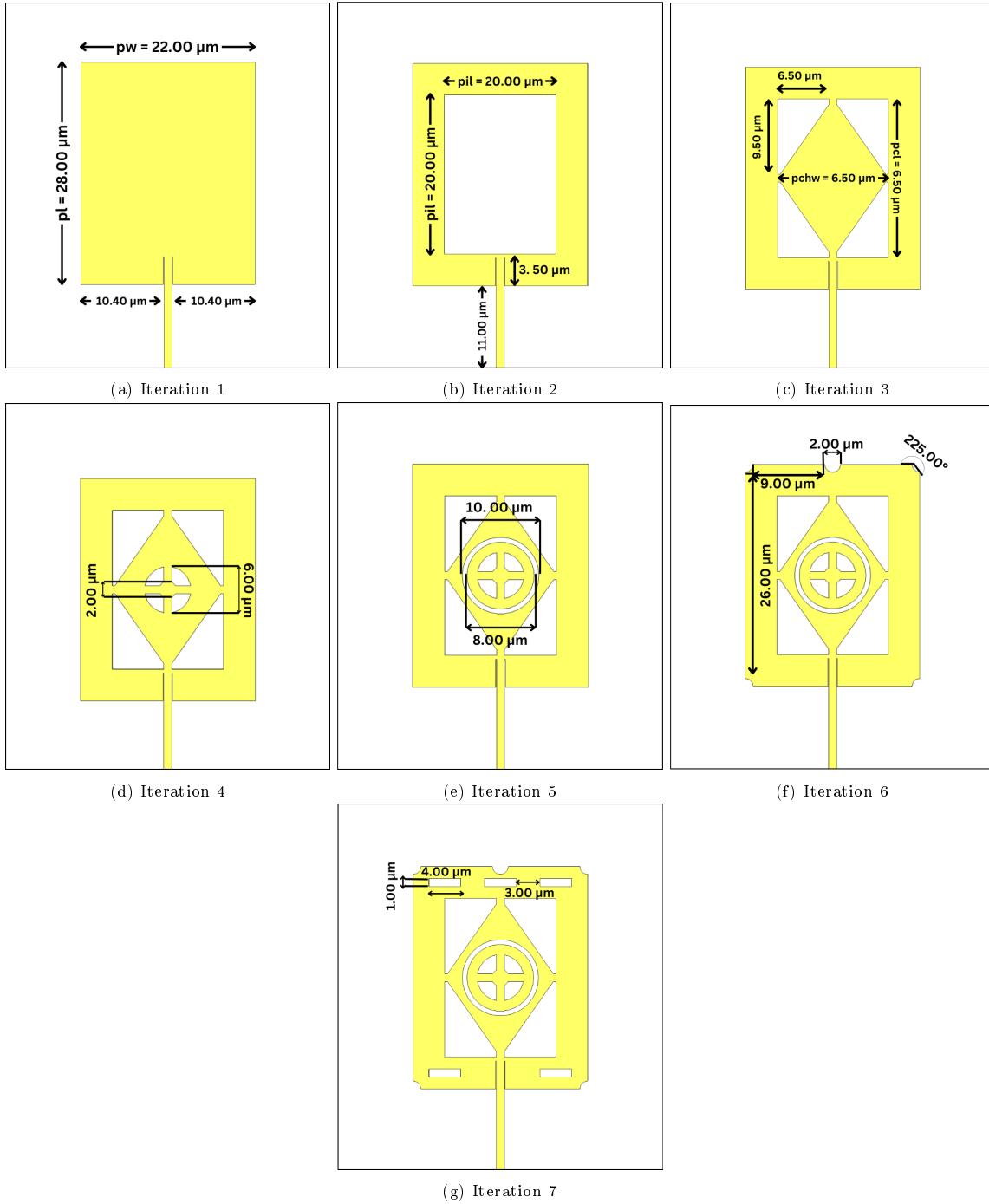


Figure 1: Advancement of antenna designs through iterative enhancements. This series of images illustrates the advancements achieved at each stage, focusing on the optimization of key performance indicators.

substrate, ensures enhanced overall performance of the antenna, making it well suited for wireless communication applications in the terahertz (THz) domain. The evolution of antenna design is divided into seven key phases, as shown in Figure 1. The following sections examine each phase of the design process in detail, from initial concept to final implementation, while highlighting important advancements and innovations in antenna technology. The following sections detail the design process, from initial concept to final implementation. Table 1 lists the geometry's parameters and values.

Table 1: Overview of Optimized Design Parameters

Design parameters	Dimension (μm)	Description
gh	0.1	Ground height
ph	0.4	Patch height
sh	3	Substrate height
sw	41	Substrate width
sl	50	Substrate length
pl	28	Exterior patch length
pw	22	Exterior patch width
pil	20	Interior patch length
piw	14	Interior patch width
pcw	1	Vertical patch width
fw	1	Feedline width
ri	1	Interior torus minor radius
ro	3	Interior torus major radius
pchw	14	Horizontal patch width
Ri	4	Exterior torus minor radius
Ro	5	Exterior torus major radius
iw	0.1	Inset width
il	3.5	Inset length

2.2. Initial Patch Structure Design

The design process commences with the symmetrical division of the rectangular patch along both the horizontal and vertical axes, as illustrated in Figure 1(a). Following this, a second rectangular patch, measuring $20 \mu\text{m} \times 14 \mu\text{m} \times 0.4 \mu\text{m}$, is incorporated within the initial patch. The inner patch is subsequently extracted from the outer patch, yielding a free rectangular region with dimensions of $20 \mu\text{m} \times 14 \mu\text{m} \times 0.4 \mu\text{m}$, as depicted in Figure 1(b).

2.3. Diamond Patch Integration and Refinement

A diamond shaped patch is subsequently positioned within the available space, with its vertices aligned to the midpoints of each side of the remaining rectangular patch. To anchor the diamond patch to the primary structure, bricks measuring $20 \mu\text{m} \times 1 \mu\text{m} \times 0.4 \mu\text{m}$ are utilized in the vertical orientation, while bricks measuring $1 \mu\text{m} \times 14 \mu\text{m} \times 0.4 \mu\text{m}$ are employed in the horizontal orientation, as illustrated in Figure 1(c).

2.4. Design Refinement with Circular Patches

The subsequent phase involves the fabrication of a torus shaped patch, characterized by an outer diameter of $3 \mu\text{m}$ and an inner diameter of $1 \mu\text{m}$, centered within the diamond shaped patch. Upon excising this torus, a cylindrical void with a height of $0.4 \mu\text{m}$ is created, as depicted in Figure 1(d). Furthermore, a second torus is generated, possessing an outer diameter of $5 \mu\text{m}$ and an inner diameter of $4 \mu\text{m}$, which is similarly removed to further refine the design, as illustrated in Figure 1(e).

2.5. Border Carving and Slot Integration

In the concluding stages of design refinement, the edges of the main frame and the central points along the width sides are meticulously shaped into arcs with a radius of $1 \mu\text{m}$, as illustrated in Figure 1(f). Furthermore, three rectangular slots, each measuring $1 \mu\text{m} \times 4 \mu\text{m} \times 0.4 \mu\text{m}$, are precisely cut into the width sides of the main frame. These slots are strategically positioned $2 \mu\text{m}$ from the edges and $1.5 \mu\text{m}$ from both the top and bottom of the frame. To maintain symmetry, corresponding rectangular slots are also carved into the opposite width side, as depicted in Figure 1(g). These modifications result in the finalized antenna design as shown in Figure 1(g).

2.6. Microstrip Feedline Configuration

The antenna design features a microstrip feedline, a prevalent technique known for its efficiency in energy transmission within planar antenna configurations, particularly in the terahertz spectrum [11]. In this arrangement, the feedline extends from a vertical connecting patch that is securely anchored to the substrate at one end, as illustrated in Figure 1(b). The configuration of the feedline is crucial, as it significantly influences both impedance matching and overall antenna performance [12]. Initially, the feedline was dimensioned at $11 \mu\text{m} \times 1 \mu\text{m} \times 0.4 \mu\text{m}$. To enhance impedance compatibility with the radiating patch, rectangular insets measuring $3.5 \mu\text{m} \times 0.1 \mu\text{m} \times 0.4 \mu\text{m}$ were deliberately removed from the edges of the feedline. This modification aimed to optimize power transfer, ensuring efficient energy delivery to the antenna. During the optimization phase, it became evident that the original feedline width of $3 \mu\text{m}$ was inadequate, resulting in poor return loss and misalignment of impedance due to insufficient compatibility with the antenna structure. This mismatch led to increased signal reflections, negatively impacting performance. Reducing the feedline width to $1 \mu\text{m}$ resulted in some improvements; however, further narrowing the width caused a decline in performance, suggesting the presence of an optimal width for effective operation. This highlights the vital role that precise feedline dimensions play in achieving the desired electrical characteristics and overall functionality of the antenna.

3. Results and Performance Analysis of Antenna Design

In this study, we utilized CST Microwave Studio, which employs the finite integral technique to analyze the proposed antenna design [4]. This software is particularly suited for addressing complex electromagnetic structures with high accuracy. Simulations were performed under open boundary conditions to effectively simulate a free space environment. To enhance precision in our results, we implemented a hexahedral mesh with adaptive refinement [10].

The performance of various antenna designs was evaluated based on return loss, a critical parameter indicating the level of signal reflection due to impedance mismatches [10]. A higher return loss signifies improved impedance matching and reduced reflected power. As depicted in Figure 2, the return loss across different design iterations was analyzed. A detailed analysis of return loss across iterations reveals the significant impact of slotting on performance metrics. Figure 2(a) illustrates the return loss for the basic rectangular microstrip patch antenna, which experienced considerable improvements after incorporating slots. The introduction of an additional rectangular slot from the main structure resulted in a noticeable reduction in return loss. Subsequent modifications involving slot incorporation led to further enhancements, culminating in the final design iteration achieving an optimal return loss of -48.974 dB at a resonant frequency of 7.713 THz . This remarkably low return loss reflects exceptional impedance matching, ensuring efficient transmission of incident power with minimal reflection.

In addition to return loss, we assessed the antenna's gain, another vital performance metric. The antenna demonstrated a maximum gain of 8.102 dB at the resonant frequency of 7.713 THz , as illustrated in Figure 3(b). A higher gain indicates an improved ability to direct radiated energy toward the intended target. Furthermore, the measured bandwidth of the antenna was 848.934 GHz , demonstrating its capacity for efficient operation across a wide frequency range.

To further optimize the antenna design process, we integrated machine learning techniques, particularly artificial neural networks (ANNs), to predict critical performance metrics, such as return loss and gain, based on varying design parameters. This machine learning driven approach significantly accelerated the design iteration process by reducing reliance on exhaustive simulations and manual adjustments. By training the

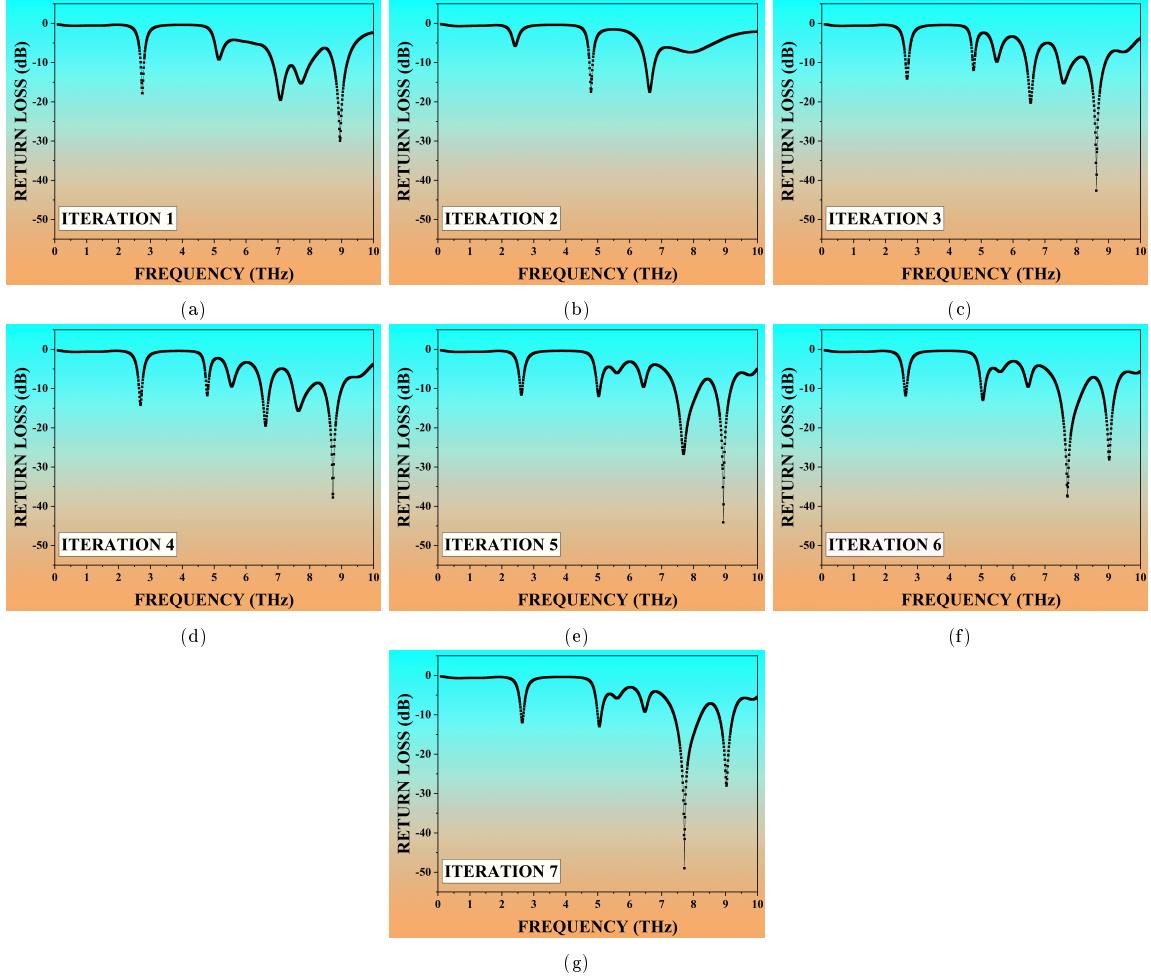


Figure 2: Return Loss Variation Across Design Iterations. The figure illustrates the progressive improvement in return loss with each design iteration, highlighting the enhancements made in minimizing signal reflection and optimizing impedance matching.

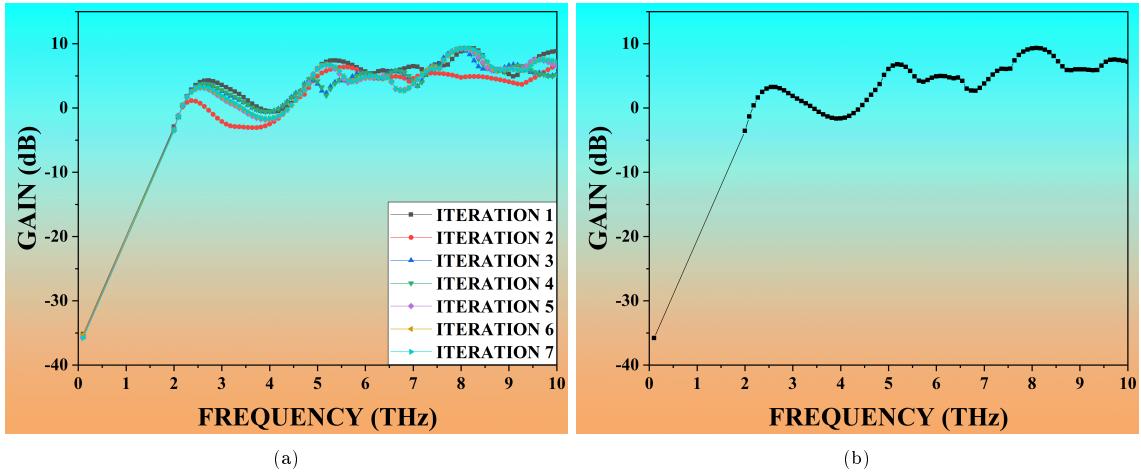


Figure 3: (a) Antenna gain progression over iterations. This figure highlights the evolution of antenna gain across design iterations. (b) The optimal design iteration achieved a maximum gain of 8.102 dB at a resonant frequency of 7.713 THz, demonstrating enhanced directivity and radiation efficiency.

ANN on a large dataset generated through CST simulations, the model was able to identify patterns in the data and predict optimal configurations for improved antenna performance.

Machine learning provided key insights into parameter interactions that traditional methods may overlook. For instance, the ANN appeared to uncover intricate dependencies between certain design parameters and performance metrics that were not readily noticeable during the initial simulation phase. This predictive capability not only reduced the overall design time but also improved the accuracy and reliability of the final antenna configuration. The integration of computational intelligence through machine learning allowed us to explore a larger design space, leading to the final design iteration, which achieved both optimal return loss and maximum gain.

Incorporating machine learning into the design process represents a significant advancement for antenna engineering. By leveraging its predictive power, we were able to streamline the development process, minimizing the need for time intensive simulations and accelerating the identification of optimal configurations. This approach holds promise for future research, where further machine learning techniques, such as ensemble methods or reinforcement learning, could be applied to optimize increasingly complex antenna designs and operational environments. Machine learning's potential to improve antenna performance metrics, such as gain and return loss, suggests a promising direction for future work in the design and optimization of advanced antenna systems.

4. Machine Learning Driven Antenna Design Optimization

Machine Learning (ML), a sub field of artificial intelligence (AI), is rapidly emerging as a key technology in a variety of professions [7]. It consists of the development of algorithmic and statistical frameworks that enable computers to carry out tasks using patterns and inference rather than explicit instructions [7]. The ability of machine learning to derive information from data and evolve superiorly over time has made it a vital tool for designing and optimizing complex systems, like antennas. Traditional design and optimization methods often involve iterative processes that can be time consuming and computationally expensive. But by allowing quick assessment in significant design areas, machine learning may completely reinvent this method of evaluation.

ML algorithms can be trained to understand the correlations between performance metrics and design factors using methods like parametric sweeps. After training, these models are capable of accurately predicting how new designs would perform, significantly lowering the need for laborious simulations. The upcoming era of AI and ML will further enhance the capabilities and complexities of antenna designs. ML models will get even more precise and quick as computational capability advances and more complex and refined algorithms are created. These models can be adopted with other AI technologies to create completely automated design workflows. Futuristic antennas will be more sophisticated and versatile as a result. Antenna optimization will be significantly impacted by AI and ML technologies as they develop, opening the door to enhanced and efficient communication systems. The implementation of machine learning into this proposed antenna design marks an important milestone in the field of antenna design.

4.1. Artificial neural network (ANN)

Artificial Neural Networks (ANNs) are composed of multilayer perceptrons (MLPs) that excel in recognizing patterns and addressing intricate problems through data driven learning [2, 8]. An ANN is structured with layers of interconnected nodes, referred to as "neurons," organized into three primary categories: an input layer, one or more hidden layers, and an output layer, as depicted in Figure 4 [2]. Each neuron processes incoming information via weighted connections, applying an activation function to produce an output that is transmitted to the subsequent layer. The strength of these connections, or weights, is refined during the training process to enhance the network's performance.

The design of an ANN begins with defining its architecture, which includes determining the number of layers, the number of neurons within each layer, and the choice of activation functions. Common activation functions, such as sigmoid, ReLU (Rectified Linear Unit), and linear, play a crucial role in enabling the network to effectively capture non-linear relationships in the data. Training an ANN focuses on minimizing a loss function that assesses the difference between predicted outputs and actual target values. This process is typically accomplished through backpropagation, wherein the error is propagated from the output layer back to the input layer, allowing for the adjustment of weights using gradient descent or other optimization

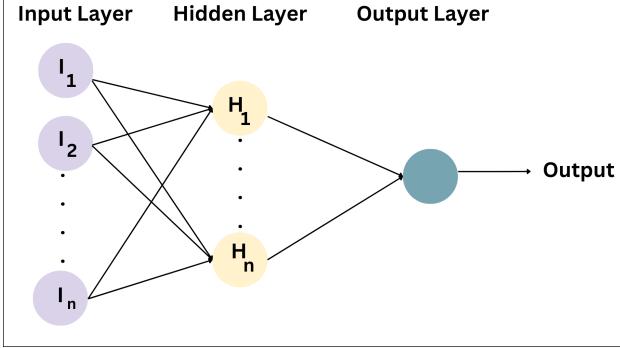


Figure 4: Schematic representation of a basic multilayer perceptrons (MLP) architecture, illustrating the interconnections between input, hidden, and output layer.

techniques [8, 13]. The goal is to identify the optimal set of weights that minimizes the loss function, thereby enhancing the accuracy of the network.

To mitigate overfitting and ensure the model generalizes effectively to new data, regularization techniques such as dropout and weight decay are commonly implemented. Furthermore, the choice of hyperparameters, including learning rate, batch size, and the number of training epochs, significantly influences the success of the training process [13]. ANNs have gained widespread adoption across various domains due to their capacity to model complex, non-linear relationships, with applications spanning from image and speech recognition to predictive modeling in engineering, such as optimizing antenna designs.

4.2. Key Equations in the Functioning of Artificial Neural Networks

Artificial neural networks represent sophisticated computational frameworks modeled after the human brain, aimed at identifying patterns and making informed decisions based on data inputs [2]. At their core, ANNs rely on a series of mathematical equations that govern their operation, training, and performance. This section explores the foundational equations that drive the functionality of ANNs, emphasizing their importance and the connections among them.

4.2.1. Activation function

A key equation in the framework of Artificial Neural Networks is the activation function, which assesses whether a neuron will be activated by evaluating the weighted sum of its input signals [2]. The activation functions incorporated in our proposed models comprise the two most widely recognized types: the Rectified Linear Unit (ReLU) and linear activation functions.

The ReLU activation function is defined mathematically as:

$$g(x) = \max(0, x) \quad (1)$$

For any input x :

- If $x > 0$, then $g(x) = x$
- If $x \leq 0$, then $g(x) = 0$

This behavior effectively allows the function to activate only for positive inputs, thereby introducing non-linearity into the model.

The linear activation function is defined mathematically as:

$$g(x) = x \quad (2)$$

In this scenario, the output is directly equal to the input, which facilitates an unaltered mapping of values. This characteristic is particularly beneficial in tasks requiring a linear relationship between inputs and outputs, such as regression analysis.

4.2.2. Output Mechanism

The output of a neuron in an artificial neural network is represented by the equation:

$$\hat{y} = g \left(\sum_{i=1}^n W_i X_i + b \right) \quad (3)$$

where \hat{y} is the predicted output, X is the input vector, W is the weight vector, b is the bias value and g is the activation function.

The expression WX represents the weighted sum of the input features, with the weight vector W indicating the significance of each feature within the input vector X [2]. This mechanism enables the network to emphasize specific inputs during the training phase. Additionally, the incorporation of the bias term b provides further flexibility, allowing the model to adjust the activation function. This adjustment is essential for accurately modeling data that may not intersect the origin.

4.2.3. Loss Function

The performance of an artificial neural network is commonly assessed through a loss function, which evaluates the disparity between the predicted outputs generated by the model and the actual target values. This loss function serves as a critical metric, providing a quantitative measure of the model's accuracy and guiding the optimization process during training. By minimizing the loss, the network improves its predictions, thereby enhancing its overall effectiveness in learning from the data.

The Loss function to calculate the Mean Squared Error (*MSE*) is given below:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (t_i - \hat{y}_i)^2 \quad (4)$$

where n is the total number of data points, t_i is the actual output and \hat{y}_i is the predicted output.

4.2.4. Backpropagation

Backpropagation is a fundamental algorithm used in the training of artificial neural networks facilitating the learning process by systematically adjusting weights and biases to reduce the discrepancy between predicted outputs and actual target values [8, 13]. This algorithm, when combined with gradient computation and parameter updates, forms the foundational framework for learning in artificial neural networks. The backpropagation process involves two primary phases: the forward pass and the backward pass.

1. Forward Pass: In the forward pass, input data is transmitted through the network sequentially, layer by layer, culminating in the generation of an output. This output is then evaluated against the target value using a loss function, which quantifies the error incurred.
2. Backward Pass: During the backward pass, the algorithm computes the gradient of the loss function with respect to each weight in the network by employing the chain rule of calculus. These gradients provide critical information about how the loss will vary in response to infinitesimal adjustments in the weights, thereby guiding the optimization process to enhance model performance.

The gradient of the loss function L with respect to weights W and bias b is represented as follows:

$$\frac{\partial L}{\partial W} = \frac{1}{n} \sum_{i=1}^n (t_i - \hat{y}_i) \quad (5)$$

$$\frac{\partial L}{\partial b} = \frac{1}{n} \sum_{i=1}^n (t_i - \hat{y}_i) \quad (6)$$

where $\frac{\partial L}{\partial W}$ represents the gradient of the loss function L with respect to the weight vector W and $\frac{\partial L}{\partial b}$ represents the gradient of the loss function with respect to the bias b .

Upon calculating the gradients, the weight vector W and the bias b are adjusted using gradient descent. The corresponding update rules are formulated as follows:

$$W \leftarrow W - \eta \frac{\partial L}{\partial W} \quad (7)$$

$$b \leftarrow b - \eta \frac{\partial L}{\partial b} \quad (8)$$

Here, η denotes the learning rate, which regulates the magnitude of the parameter updates. The selection of an appropriate learning rate is critical for effective training; a rate that is too small can impede convergence, while a rate that is excessively large may lead to instability or divergence.

4.3. Data Collection

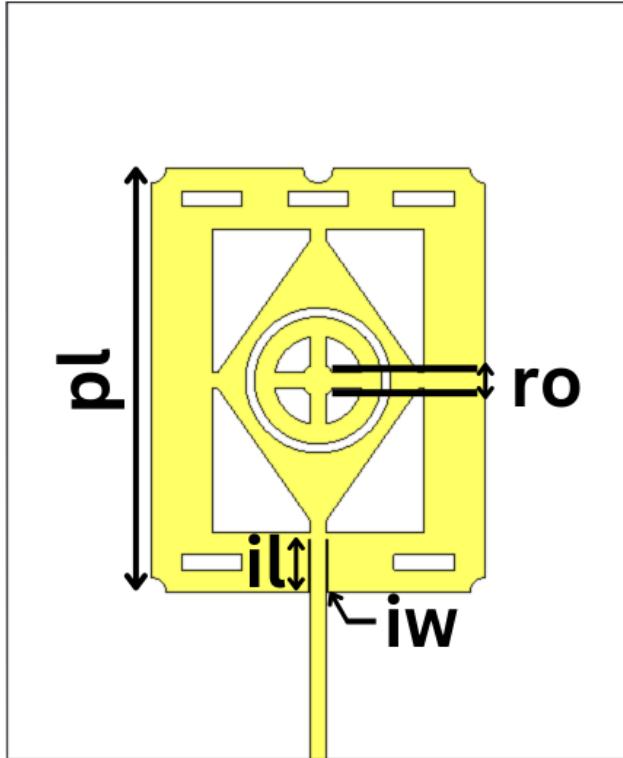


Figure 5: Iterated Design Parameters for Data Generation. This figure presents the parameters, il , iw , pl , and ro , used to assess their effects on target performance metrics.

To train the proposed machine learning model, we carefully compiled a dataset comprising 1296 design permutations of the optimized antenna. The parameters selected for variation consist of the inset length (il), inset width (iw), the length of the exterior rectangular patch (pl), and the larger radius of the interior torus-shaped patch (ro), as demonstrated in Figure 5. Each of these parameters plays a critical role in influencing the overall design and performance of the antenna, allowing for a comprehensive analysis of their impact on the system's effectiveness. A comprehensive parametric sweep was conducted using CST Microwave Studio to generate these permutations, ensuring a broad exploration of the design space. The range of values for each parameter, as detailed in Table 2, was systematically altered to produce 1296 unique outcomes, each representing a distinct configuration of the antenna. The corresponding performance metrics; resonating frequency ($freq$), return loss (S_{11}), and gain, were recorded as target variables, with the data being systematically extracted and stored in a CSV file format. Prior to model training, the dataset underwent rigorous preprocessing to ensure consistency and accuracy. This included error checking to remove any anomalies or missing values that could potentially degrade model performance. The dataset was then partitioned into two segments: an 85% training set (1102 data points) and a 15% validation set (194

data points). This segmentation strategy was implemented to facilitate robust model training and reliable performance evaluation.

Given the diversity in data types, ranging from positive and negative values to integers and floating point numbers, feature scaling was employed to standardize the dataset. Specifically, MinMax scaling was utilized, normalizing each parameter to a range between 0 and 1 [14]. This preprocessing step was crucial to ensuring that all parameters contributed equally to the model, thereby enhancing the accuracy and reliability of predictions while maintaining proportional relationships among attributes.

The equation for MinMax scaling transformation is given below [14]:

$$X_{\text{scaled}} = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \quad (9)$$

where X_{scaled} is the scaled value of the feature X_i , X_i is the original value of the feature that is being scaled, X_{\min} is the minimum value of the feature across the dataset and X_{\max} is the maximum value of the feature across the dataset.

Table 2: Iterated Parameter Ranges for Model Training

Sq.No.	Parameter	Range (μm)	Step width (μm)	Data set
1	il	3.0 - 3.5	0.1	6
2	iw	0.1 - 0.6	0.1	6
3	pl	28.0 - 35.0	1.3	6
4	ro	3.0 - 3.5	0.1	6

4.4. Data Visualization

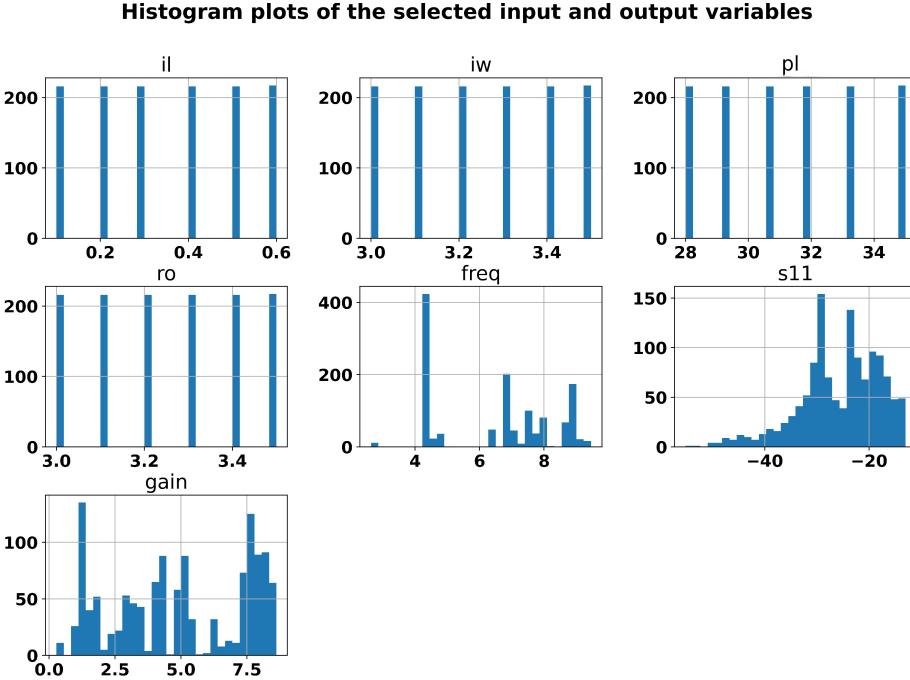


Figure 6: Histogram plots of the selected input and output variables. The input parameters (il, iw, pl, ro) are uniformly distributed, while output parameters (freq, S_{11} , gain) exhibit more distinct or skewed distributions, reflecting the underlying system behavior and performance measures.

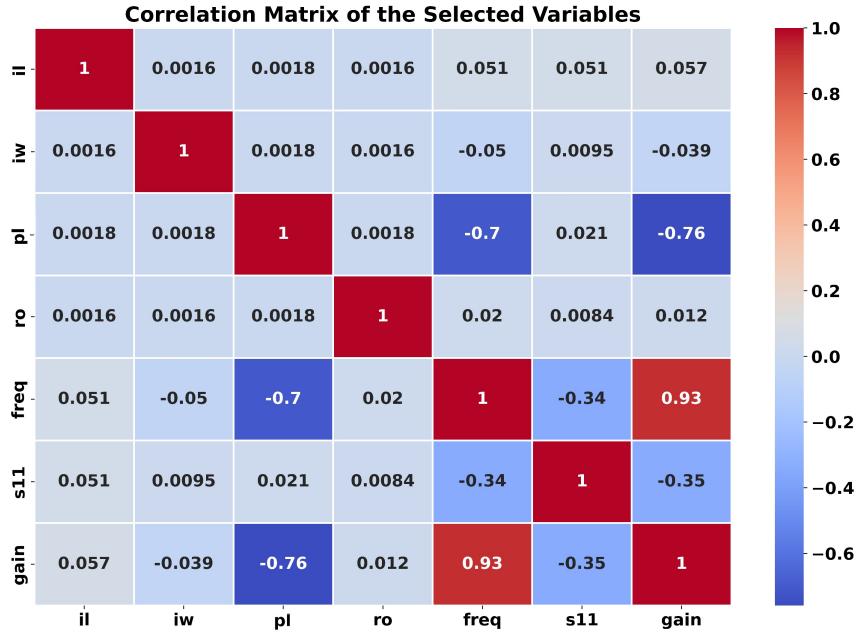


Figure 7: The correlation matrix of the selected variables, showing the strength and direction of relationships between input and output features. Strong correlations are marked by darker shades of red (positive) or blue (negative), indicating direct or inverse relationships, respectively.

Data visualization is essential for interpreting complex datasets in antenna design, where understanding patterns and relationships among variables is critical. By converting raw data into graphical formats, it becomes easier to identify trends that guide decision making and optimize model performance. In this study, we use histograms and correlation matrices to assess the distribution of input parameters and output features, crucial for evaluating how well the data covers the design space. Histograms offer clear insights into variable spread and clustering, while correlation matrices reveal the strength of relationships between variables, informing key balances in antenna performance and guiding model optimization. These tools were specifically chosen for their ability to provide both detailed and relational insights into the data, enhancing the model's predictive capabilities for antenna design optimization.

Figure 6 illustrates the histograms of the selected input parameters (il, iw, pl, ro) and output features (freq, S_{11} , gain), shedding light on their distribution across the dataset. The distribution of these variables is a crucial aspect to consider when evaluating model training and performance, as it reveals how well the dataset represents the potential variations in both input and output spaces. The input variables (il, iw, pl, and ro) show evenly distributed histograms, suggesting a nearly uniform spacing between their values. This even distribution is indicative of a well balanced dataset, providing comprehensive coverage across the ranges of the input variables. Such balance is vital in ensuring that the model encounters a wide variety of scenarios during training, allowing it to generalize effectively to different configurations during the testing and deployment stages.

On the other hand, the output feature frequency displays a more concentrated distribution, with values clustering around specific points, particularly at approximately 4, 6, and 8 units. This indicates that the dataset focuses on certain critical frequency bands rather than covering a continuous frequency spectrum. Such clustering suggests a deliberate selection of specific frequency points for analysis, possibly due to their significance in the design and operational contexts of the antenna system. The concentration of frequency values around these bands could also affect the model's predictive capabilities, allowing it to perform more accurately within these focused regions while potentially being less accurate outside them. This characteristic is typical in many engineering applications, where particular frequency bands hold more importance due to

system design constraints or operational requirements.

In contrast, the S_{11} feature exhibits a skewed distribution, with a higher concentration of values between -50 and -20. This skewness toward more negative values is advantageous in many electromagnetic systems, as more negative S_{11} values indicate better reflection coefficient performance, which corresponds to less power being reflected back from the antenna. Such a performance metric is essential in many antenna designs where minimizing power reflection is critical for efficient operation. The skewed distribution of S_{11} reflects the challenges in achieving highly negative values, as optimal return loss is difficult to realize in practical systems. This non-uniform distribution provides important context for understanding how the model might perform under different conditions, especially in scenarios where S_{11} values are more or less ideal.

The gain variable shows a broader distribution, with peaks at around 2, 4, and 8 units, indicating that the dataset captures a wide range of performance outcomes. This broader distribution reflects the diversity of system behaviors, covering both lower and higher performance scenarios. The frequent occurrence of higher gain values suggests that the dataset is designed to emphasize high performance outcomes, which are often desirable in the design of antenna systems. This variety in the gain values ensures that the model is exposed to a range of performance levels during training, allowing it to learn the relationships between the input parameters and the desired high performance outcomes. This diversity within the dataset also supports a more robust training process, as the model learns to handle a wider array of performance regimes.

Overall, the histogram analysis provides valuable insights into the dataset's structure. While the input variables demonstrate a well balanced, uniform distribution, the output variables show more structured or skewed patterns, reflecting real world performance constraints and optimization goals. The clustering of frequency around specific points, the skewness of S_{11} , and the broader distribution of gain highlight key aspects of the data that must be accounted for when interpreting the model's performance and predictive capabilities. These distributions directly impact how the model generalizes across different regions of the input and output space, particularly when predicting performance metrics that are central to antenna design and optimization.

In addition to the histograms, Figure 7 presents the correlation matrix of the selected input and output variables, offering further insights into the relationships between them. This matrix is a powerful tool for analyzing how changes in one variable may affect others, allowing for a deeper understanding of the dependencies and interactions within the dataset. The input variables (il, iw, pl, and ro) exhibit very weak correlations with one another, as expected, since they are likely independent design parameters. For instance, il and iw show a minimal correlation of 0.0016, while pl and ro demonstrate a similarly low correlation of 0.0018. These weak correlations suggest that changes in one input variable do not significantly impact the others, which is beneficial for antenna design processes where independent parameter variations are necessary to explore a wide range of design configurations. The independence of the input variables ensures flexibility in adjusting each parameter without influencing others, a critical aspect of efficient optimization in design workflows.

The frequency variable displays a moderate negative correlation with pl (-0.7), indicating that as the patch length (pl) increases, the frequency tends to decrease. This relationship reflects a fundamental physical dependency between these two parameters in electromagnetic system design, where larger patch lengths typically result in lower resonant frequencies. This inverse relationship is crucial in understanding how structural modifications to the antenna, such as changes in patch length, can influence its operating frequency. Additionally, frequency shows a strong positive correlation with gain (0.93), suggesting that higher frequencies are associated with higher gain. This correlation is particularly important for system performance, as achieving higher gain at higher frequencies is often a key design objective in antenna systems, enhancing signal transmission and reception capabilities.

The S_{11} feature demonstrates a weak to moderate negative correlation with both frequency (-0.34) and gain (-0.35). This suggests that as frequency and gain increase, S_{11} tends to become less negative, implying a decline in performance. Since more negative S_{11} values indicate better performance in terms of power reflection, this inverse relationship highlights the challenges of maintaining optimal return loss at higher frequencies and gains. The gain variable exhibits a very strong positive correlation with frequency (0.93), confirming that higher gain is typically achieved at higher frequencies, while better S_{11} values are more likely to occur at lower gain levels. These relationships provide critical insights for performance optimization, guiding the trade offs that need to be made between frequency, gain, and return loss during the design process.

The correlation matrix provides a comprehensive overview of the interdependencies between input and

output variables, offering key insights into how these variables interact and influence one another. The strong correlation between frequency and gain, along with the inverse relationship between pl and frequency, underscores the importance of understanding these relationships when optimizing the model and the system as a whole. These correlations will play a crucial role in guiding design decisions and improving the overall performance of the neural network model in predicting antenna behavior.

5. Model Training

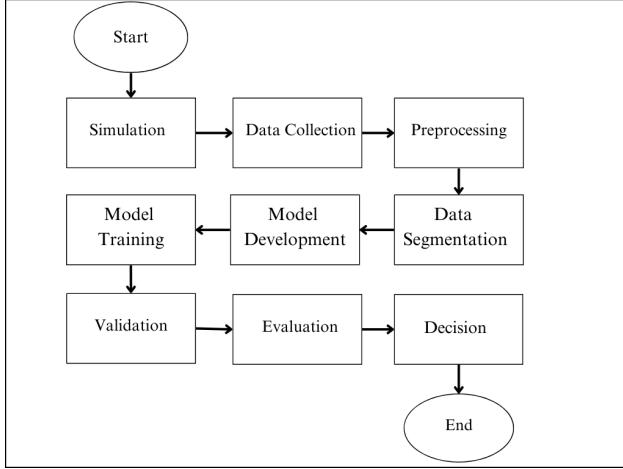


Figure 8: Workflow for the proposed neural network model development. This flowchart illustrates the key phases in the neural network development process, from simulation to evaluation, emphasizing the integration of machine learning to optimize antenna performance.

After scaling the input data, we developed a neural network model (regressor) aimed at predicting the target traits. The dataset was split into training and testing sets using Scikit-learn. The neural network was designed with four input dimensions, corresponding to the four input parameters. The architecture begins with a sequential model, featuring an initial dense layer of 64 neuron units utilizing the 'ReLU' activation function, resulting in a total of 320 parameters. The input data is further processed through a second and third dense layer, with 512 and 1024 neuron units respectively, significantly enhancing the model's learning capacity. To reduce dimensionality before the final output, a fourth dense layer comprising 512 neuron units was incorporated. The output layer, consisting of 3 neuron units corresponding to the 3 target traits, is activated using a 'Linear' function. The choice of the Rectified Linear Unit (ReLU) and linear activation functions in our models is driven by ReLU's computational efficiency and its ability to alleviate the vanishing gradient problem, while the linear activation function provides a straightforward approach for regression tasks. This combination leverages the strengths of both functions, enhancing the models' performance across diverse applications. The detailed architecture of the proposed model is illustrated in Table 3. For model validation, we initially trained the model for 100 epochs. Recognizing that ANNs generally achieve better convergence with higher epoch values, we extended the training to 1000 epochs. Figure 8 depicts the comprehensive approach followed for developing the proposed model.

6. Results and Performance Evaluation of the Proposed Neural Network Model

The effectiveness of the proposed neural network model was rigorously assessed using several critical performance metrics, including accuracy, loss, Mean Absolute Error (MAE), Mean Squared Error (MSE), and the R² score. These metrics, as detailed in Table 4, collectively provide a comprehensive evaluation of the model's predictive capabilities and its ability to generalize to previously unseen data, which is particularly vital for optimizing antenna designs.

The model's performance was evaluated on both the validation and test sets, revealing strong generalization capabilities and consistent accuracy across datasets. As shown in Table 4, on the validation set, the

Table 3: Neural Network Architecture for Predicting Antenna Performance

Layer	Type	Neuron Units	Activation Function
Input	Input Layer	4	
Dense 1	Fully Connected	64	ReLU
Dense 2	Fully Connected	512	ReLU
Dense 3	Fully Connected	1024	ReLU
Dense 4	Fully Connected	512	ReLU
Output	Fully Connected	3	Linear

model achieved an accuracy of 93.15%, indicating its ability to correctly predict approximately 93.15% of the data. The low MSE of 0.0046 on the validation set demonstrates that the model's predictions closely align with the true values, while the MAE of 0.0271 further substantiates its precision, reflecting a minimal average prediction error. Additionally, the R^2 score of 0.9272 on the validation set signifies that 92.72% of the variance in the target variable was successfully explained by the model, highlighting its ability to capture complex relationships within the dataset. These results are further illustrated in Figure 9(a), where the accuracy curve shows consistent performance across epochs during training, and Figure 9(b), which visualizes the low MSE achieved during model training. On the test set, the model's performance remained strong, although there was a slight increase in error values compared to the validation set. As detailed in Table 5, the MSE on the test set was 0.0059, a modest rise from the validation set's MSE of 0.0046, indicating a small increase in prediction error when the model encountered unseen data. This increase in MSE is typical when evaluating models on new, previously unseen data, as the model may slightly underperform due to factors such as data distribution differences. The MAE on the test set was 0.0302, which is slightly higher than the 0.0271 observed on the validation set. However, this value remains very low, suggesting that the model's predictions are still very close to the true values on the test data. The MAE metric, which represents the average absolute error in the predictions, indicates that even with the slight increase in error compared to the validation set, the model remains highly reliable for practical applications. The R^2 score on the test set was 0.8925, which is slightly lower than the validation R^2 score of 0.9272. While the slight drop in R^2 is expected when applying the model to new data, the R^2 score of 0.8925 still demonstrates that the model is able to explain approximately 89.25% of the variance in the target variable. This suggests that the model continues to capture the underlying patterns of the data effectively, even when tested on unseen examples. These results are shown in Figure 9(c) and Figure 9(d), where the performance curve on the test set mirrors the trend observed on the validation set, with a minor increase in error metrics. The figure highlights the robustness of the model, as it maintains strong performance despite the slight increase in error when transitioning from validation data to unseen test data.

The Adam optimizer was employed to dynamically adjust the learning rate during the training process, facilitating the effective minimization of the mean squared error loss function. This optimizer contributed to the model's ability to efficiently converge to an optimal solution, ensuring accurate predictions. The loss, quantified as MSE, was consistently low across both validation and test sets, underscoring the model's ability to minimize the discrepancy between predicted and observed values.

In conclusion, the neural network model demonstrated its capacity to deliver precise, reliable predictions across both the validation and test sets. The combination of high accuracy, minimal loss, and strong R^2 scores highlights the model's potential for practical implementation in optimizing the critical parameters of microstrip patch antennas. While the model exhibited a small increase in error metrics on the test set compared to the validation set, the performance remained strong, the slight differences are within expected ranges and do not undermine the model's overall predictive power. These results underscore the model's suitability for real-world applications, offering a promising tool for advancing antenna design and optimization in high-performance scenarios.

7. Comparative Analysis of Machine Learning Models for Antenna Optimization

In this section, we will systematically compare and analyze the performance of various machine learning models applied to the same dataset, using consistent scaling and a fixed train-test split. The objective is

Table 4: Performance Metrics for the Neural Network on the Validation Set

Sq.No.	Metric	Value
1	Accuracy	0.9315
2	Mean Squared Error (MSE)	0.0046
3	Mean Absolute Error (MAE)	0.0271
4	R ² score	0.9272

Table 5: Performance Metrics for the Neural Network on the Test Set

Sq.No.	Metric	Value
1	Mean Squared Error (MSE)	0.0059
2	Mean Absolute Error (MAE)	0.0302
3	R ² score	0.8925
4	Explained Variance Score	0.8999

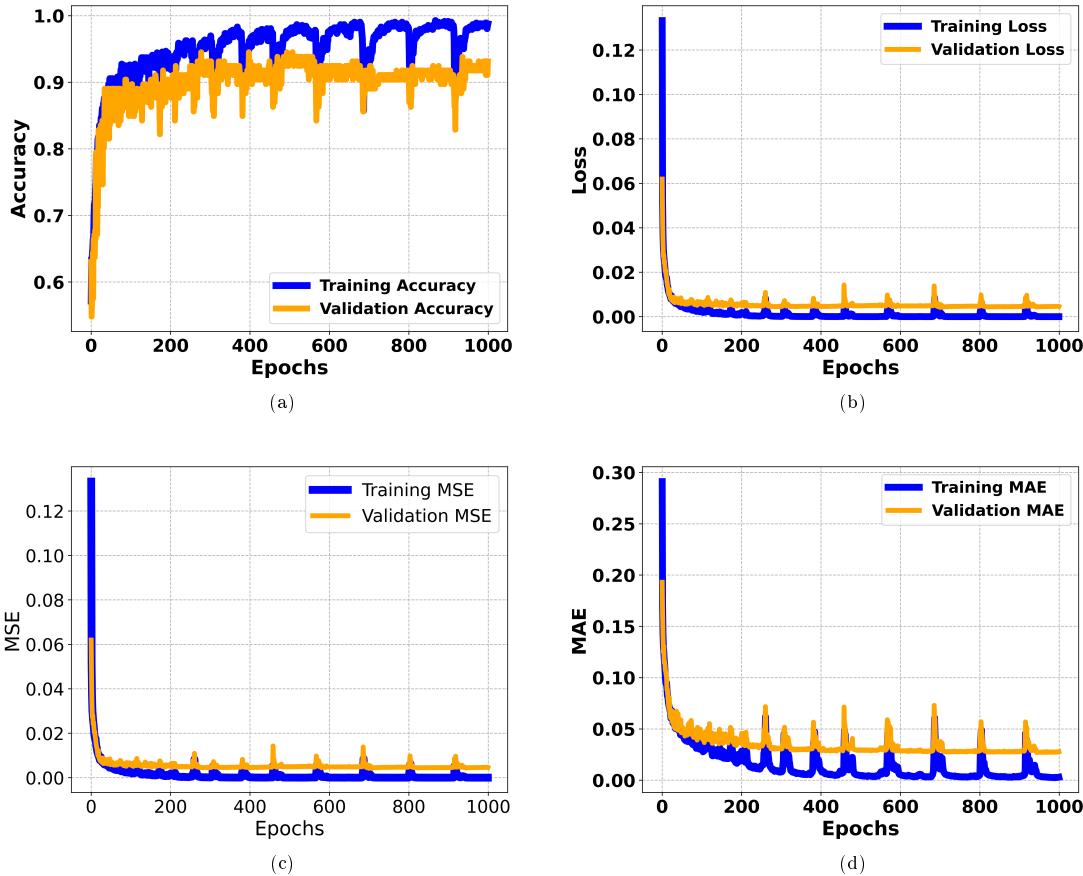


Figure 9: Model Training and Validation Metrics Over 1000 Epochs. This figure shows the evolution of key performance indicators, including (a) Accuracy, (b) Loss, (c) Mean Squared Error (MSE), and (d) Mean Absolute Error (MAE), providing a comprehensive view of model learning and validation phases.

to understand how different algorithms handle the data, their predictive accuracy, and their suitability for the problem at hand. This analysis will provide insights into the strengths and limitations of each model, ultimately guiding the selection of the most appropriate method for optimizing the microstrip patch antennas.

7.1. Dummy Model

The Dummy model is a basic benchmark in machine learning, often used to establish a minimum performance threshold. It generates predictions based on simple rules, such as consistently predicting the mean or median of the training set. Though it does not aim to provide accurate predictions, its primary role is to serve as a reference point for evaluating the added value of more advanced models. By comparing complex algorithms to this baseline, researchers can determine whether improvements are genuine or simply due to random variations.

Figure 10(a) and Figure 10(b) present scatter plots comparing predicted and actual values for the neural network model and a dummy regressor, respectively. Each plot includes a dashed line representing the line of perfect prediction, where the predicted values would ideally lie if they were perfectly aligned with the actual values.

In Figure 10(b), we see the limitations of the dummy regressor, which clusters predictions around specific values with no apparent correlation to the actual values. Table 6 further quantifies this, showing a Mean Squared Error (MSE) of 0.0618 and a Mean Absolute Error (MAE) of 0.2102, both indicating substantial deviations from the true values. The R^2 score of -0.0011 demonstrates its inability to explain variance, while the Explained Variance Score of 0.000 confirms its ineffectiveness in capturing any meaningful information from the dataset. These metrics emphasize the Dummy model's role as a foundational benchmark in predictive modeling, highlighting the necessity for more advanced techniques to achieve meaningful predictive accuracy and performance. In contrast, Figure 10(a) illustrates that the predictions made by the neural network model closely follow the line of perfect prediction, indicating a strong correlation between predicted and actual values. This alignment suggests that the neural network effectively captures the underlying patterns in the data.

Together, these results demonstrate the superiority of the neural network model over the dummy regressor. The significant improvements in error rates and R^2 score validate the pursuit of sophisticated algorithms to address the complexities of real-world datasets. By establishing a baseline with the Dummy model, we confirm that the neural network offers genuine predictive improvements, capturing complex relationships within the data and achieving a high level of accuracy and variance explanation.

Table 6: Performance Metrics for the Dummy Model

Sq.No.	Metric	Value
1	Mean Squared Error (MSE)	0.0618
2	Mean Absolute Error (MAE)	0.2102
3	R^2 score	-0.0011
4	Explained Variance Score	0.0000

7.2. Linear Regression

The Linear Regression is a cornerstone of statistical modeling and machine learning, renowned for its straightforwardness and interpretability [5]. By assuming a linear relationship between input features and the target variable, it fits a linear equation to the data, minimizing the differences between predicted and observed values through the Ordinary Least Squares (OLS) method [5]. This simplicity makes the linear regression an appealing choice for initial data exploration, allowing researchers to grasp how changes in input variables affect predictions. However, its inherent limitations become apparent in complex datasets, where non-linear relationships and intricate interactions may yield misleading outcomes [5]. Therefore, while it serves as a valuable tool for understanding fundamental relationships, there is a pressing need to explore more sophisticated modeling techniques to enhance predictive accuracy in real world applications.

Table 7 presents the key performance metrics for the linear regression, providing valuable insights into its predictive capabilities. The Mean Squared Error (MSE) is reported at 0.0296, indicating a relatively low

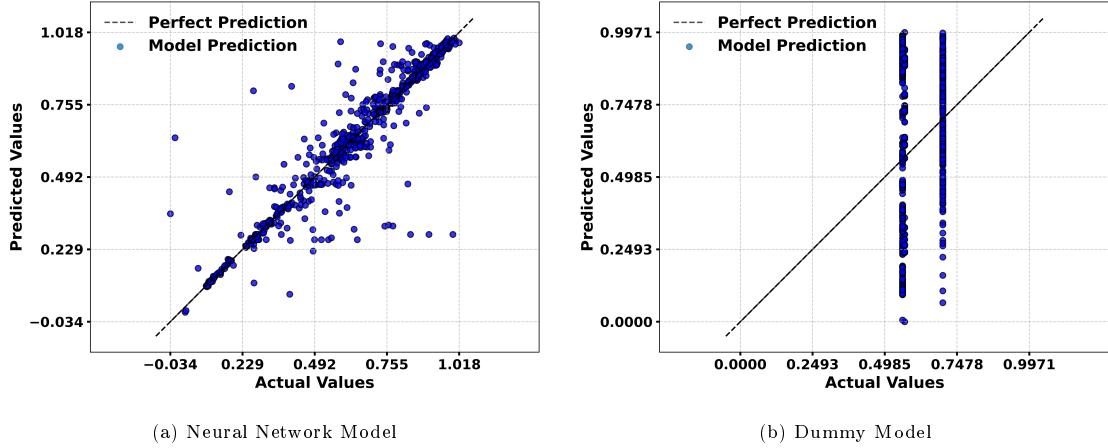


Figure 10: Comparison of Actual and Predicted Values. The figure contrasts the predictions made by (a) the Neural Network model and (b) the Dummy Model against the actual values, illustrating the performance of each approach.

level of error in the model’s predictions. The Mean Absolute Error (MAE) is recorded at 0.1348, reflecting the average magnitude of the errors in the predictions, suggesting that there are some deviations from the actual values. The R^2 score of 0.4000 indicates that approximately 40% of the variance in the target variable can be explained by the Linear model. While this score demonstrates some level of explanatory power, it also suggests significant room for improvement, as the model does not capture all relevant variability within the dataset. Additionally, the Explained Variance Score of 0.4009 corroborates this finding, further highlighting the limitations of the model in effectively capturing the underlying data patterns.

The scatter plot in Figure 11 visually underscores these metrics by comparing the predicted values against actual values. Notably, the scattered points, especially in the mid-range values, indicate that the linear model struggles with accurately capturing the full complexity of the data, likely due to non-linear relationships or intricate interactions it cannot model. These visual and numerical insights together underscore the need to explore more advanced models that can better handle non-linear relationships, ultimately enabling more accurate and reliable predictions.

These metrics underscore the necessity for considering more advanced modeling techniques when faced with complex datasets exhibiting non-linear relationships. While the linear regression model serves as a useful initial tool, recognizing its limitations is crucial in the pursuit of achieving more accurate and reliable predictions.

Table 7: Performance Metrics for the Linear Regression Model

Sq.No.	Metric	Value
1	Mean Squared Error (MSE)	0.0296
2	Mean Absolute Error (MAE)	0.1348
3	R^2 score	0.4000
4	Explained Variance Score	0.4009

7.3. Ridge Regression

Ridge Regression extends the principles of linear regression by incorporating a regularization term into the loss function to enhance model stability and performance. This adjustment addresses common issues in predictive modeling, such as overfitting and multicollinearity, by penalizing large coefficients [5]. However, as illustrated in Figure 12, the model’s predicted values display a substantial deviation from the actual values, with the majority of points clustering far from the line of perfect prediction. This visual misalignment underscores the limitations of the Ridge model in this context, where it fails to capture the true underlying patterns of the data effectively.

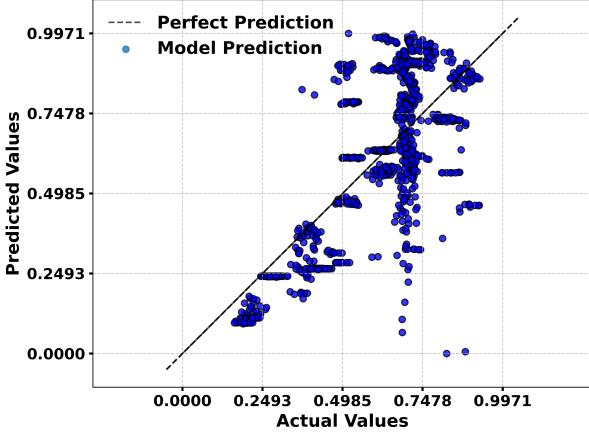


Figure 11: Linear Regression Model. This figure shows the comparison of actual versus predicted values generated by the linear model, illustrating its performance and predictive accuracy.

Table 8 presents the key performance metrics for the ridge regression model, further highlighting its significant shortcomings. The Mean Squared Error (MSE) is recorded at 252.2769, indicating substantial error in predictions, while the Mean Absolute Error (MAE) of 12.2441 reflects considerable deviations from actual values. The R^2 score of -8595.4128 and an Explained Variance Score of -19.6920 reveal a profound inadequacy in explaining the variance of the target variable, categorizing Ridge Regression as the worst performing model.

These metrics indicate the limitations of the ridge regression approach in this context. They underscore the need to reassess the model's suitability and suggest exploring alternative methods that may provide improved predictive accuracy and performance.

Table 8: Performance Metrics for the Ridge Regression Model

Sq.No.	Metric	Value
1	Mean Squared Error (MSE)	252.2769
2	Mean Absolute Error (MAE)	12.2441
3	R^2 score	-8595.4128
4	Explained Variance Score	-19.6920

7.4. Bayesian Ridge

Bayesian Ridge Regression represents a sophisticated enhancement of linear regression, incorporating probabilistic methodologies to better model variable relationships. This approach differs from traditional linear regression by assigning prior distributions to model coefficients, enabling the estimation of coefficients alongside the quantification of uncertainty associated with these estimates. By merging Bayesian inference with ridge regression principles, the model offers a detailed understanding of the variability in coefficient values, which is particularly beneficial in contexts with noisy data or outliers.

Table 9 summarizes the key performance metrics for the bayesian ridge regression model, highlighting its predictive capabilities. The Mean Squared Error (MSE) is recorded at 0.0295, indicating a relatively low level of predictive error, while the Mean Absolute Error (MAE) is measured at 0.1345, representing the average magnitude of errors in the predictions. The R^2 score of 0.4042 signifies that approximately 40.42% of the variance in the target variable can be explained by this model. Although this reflects some level of explanatory power, it also suggests room for improvement, as the model does not fully capture the dataset's variability. The Explained Variance Score of 0.4050 further emphasizes these challenges in accurately modeling underlying data patterns.

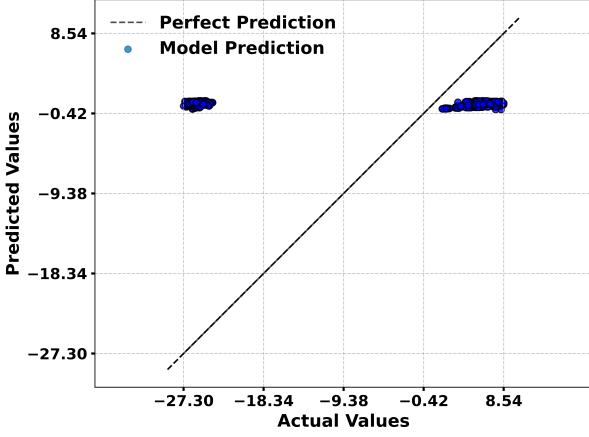


Figure 12: Ridge Regression Model. This figure shows the comparison of actual versus predicted values generated by the Ridge Regression model, illustrating its performance and predictive accuracy.

The scatter plot in Figure 13 illustrates the predictive performance of the Bayesian Ridge Regression model by comparing actual values with predicted values. The dispersion of points around this line provides a visual representation of the model's accuracy and its ability to capture relationships within the dataset. The spread of points around the perfect prediction line reveals that, while bayesian ridge regression enhances traditional linear regression by incorporating probabilistic methodologies, it still encounters challenges in perfectly matching predictions to actual values. Notably, the points exhibit clustering at certain values, which suggests limitations in capturing complex relationships or non-linear patterns within the data. However, the Bayesian approach offers an advantage by providing a measure of uncertainty for the predictions, making it valuable in contexts with inherent noise or outliers, where traditional models may struggle.

These metrics illustrate that bayesian ridge regression, while offering a robust framework for quantifying uncertainty, achieves performance levels comparable to traditional linear regression models. This finding underscores the necessity for further refinement and exploration of alternative methodologies to enhance predictive accuracy and overall performance.

Table 9: Performance Metrics for the Bayesian Ridge Model

Sq.No.	Metric	Value
1	Mean Squared Error (MSE)	0.0295
2	Mean Absolute Error (MAE)	0.1345
3	R ² score	0.4042
4	Explained Variance Score	0.4050

7.5. SVM linear kernel

Support Vector Machines (SVM) utilizing a Linear Kernel provide a fundamental approach for both classification and regression tasks, particularly effective in high dimensional datasets. The core objective of a linear kernel SVM is to find the optimal hyperplane that separates data points into distinct classes or fits data in regression scenarios, maximizing the margin between classes [5, 15]. This characteristic enhances the model's generalization capability, enabling it to perform effectively on unseen data.

Table 10 outlines the performance metrics for the SVM with a linear kernel, shedding light on its predictive capabilities. The Mean Squared Error (MSE) is reported at 0.0293, indicating a relatively low predictive error, while the Mean Absolute Error (MAE) stands at 0.1328, reflecting the average error magnitude in predictions. The R² score of 0.4050 reveals that about 40.50% of the variance in the target variable is explained by this model, demonstrating moderate explanatory power. Additionally, the Explained Variance

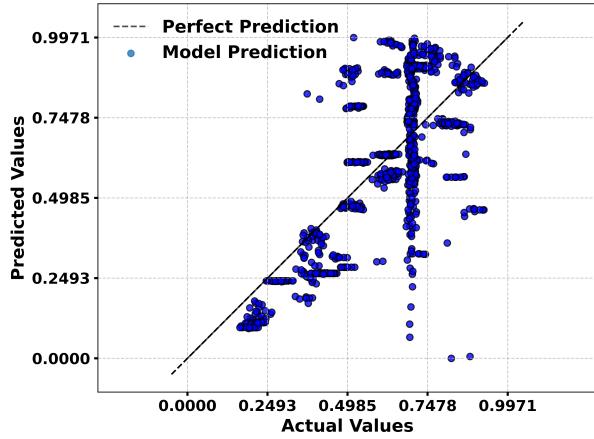


Figure 13: Bayesian Ridge Model. This figure shows the comparison of actual versus predicted values generated by the Bayesian Ridge model, illustrating its performance and predictive accuracy.

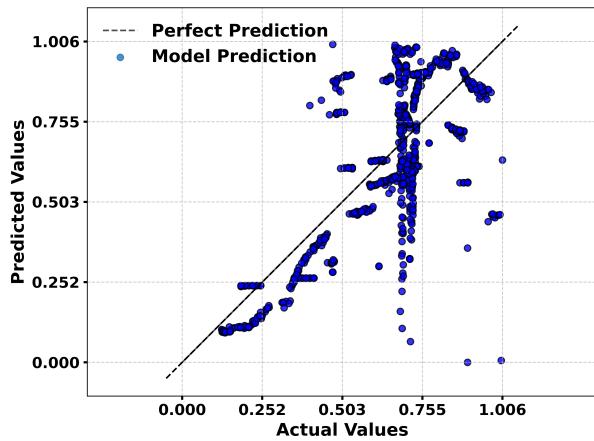


Figure 14: SVM Linear Model. This figure shows the comparison of actual versus predicted values generated by the SVM with a linear kernel, illustrating its performance and predictive accuracy.

Score of 0.4076 reinforces the notion that, while the model captures some underlying relationships, it falls short of fully accounting for dataset variability.

The scatter plot in Figure 14 visualizes the performance of the Support Vector Machine (SVM) model with a linear kernel by comparing actual and predicted values. The distribution of blue points around this line indicates the SVM's ability to approximate the target values. Observing the spread of points around the perfect prediction line, it becomes evident that the linear kernel SVM captures some underlying relationships but still encounters difficulties in fully aligning with the actual values. The moderate dispersion of points suggests that while the model performs reasonably well, it does not entirely account for all the variability in the dataset, particularly in regions where actual values are higher. This limitation is consistent with the linear kernel SVM's focus on maximizing the margin between data points, which may fall short in datasets with complex patterns or non-linear relationships.

These metrics confirm that the linear kernel SVM serves as a straightforward tool in predictive modeling, yet they highlight the necessity for more advanced techniques when addressing datasets with complex structures. This necessity underscores the importance of exploring alternative methodologies to enhance predictive performance and effectively capture intricate data patterns.

Table 10: Performance Metrics for the SVM Linear Model

Sq.No.	Metric	Value
1	Mean Squared Error (MSE)	0.0293
2	Mean Absolute Error (MAE)	0.1328
3	R ² score	0.4050
4	Explained Variance Score	0.4076

7.6. SVM polynomial kernel

The Polynomial Kernel Support Vector Machine (SVM) advances the traditional linear SVM framework, effectively modeling complex data relationships by addressing the limitations posed by non-linear patterns [12, 15]. While linear SVMs excel in straightforward scenarios, they struggle with intricate datasets where relationships are not simply linear. The polynomial kernel SVM transforms the input space into a higher dimensional domain, enabling the representation of non-linear relationships as linear ones, thus enhancing the model's ability to identify complex interactions among features [12, 15].

Table 11 provides the key performance metrics for the SVM with a polynomial kernel, illustrating its predictive capabilities. The Mean Squared Error (MSE) is reported at 0.0198, indicating a relatively low level of predictive error, while the Mean Absolute Error (MAE) is measured at 0.1060, reflecting the average magnitude of errors in predictions. The R² score of 0.6082 shows that approximately 60.82% of the variance in the target variable can be explained by this model, indicating a significant degree of explanatory power. The Explained Variance Score of 0.6119 further corroborates this effectiveness, highlighting the model's ability to account for variability in the dataset.

The scatter plot in Figure 15 offers a visual comparison of actual versus predicted values for the polynomial kernel SVM model, providing insight into its predictive performance. In this plot, it is observed that a clustering of points close to the line, indicating a reasonable correlation between actual and predicted values. However, there is also a noticeable spread, particularly at certain ranges, where the model predictions diverge from the actual values. This spread reflects some of the non-linear complexity that the polynomial kernel aims to capture, albeit with limitations.

These metrics confirm that the polynomial kernel SVM provides a robust framework for capturing complex data patterns, establishing its value in predictive modeling. This method, while effective, still emphasizes the need for careful parameter selection to avoid pitfalls like overfitting, pointing to the potential for further refinement and exploration of alternative methodologies.

7.7. SVM RBF kernel

The Radial Basis Function (RBF) Kernel Support Vector Machine (SVM) is a versatile model adept at tackling complex, non-linear relationships in data. By transforming input data into an infinite dimensional

Table 11: Performance Metrics for the SVM Polynomial Kernel

Sq.No.	Metric	Value
1	Mean Squared Error (MSE)	0.0198
2	Mean Absolute Error (MAE)	0.1060
3	R ² score	0.6082
4	Explained Variance Score	0.6119

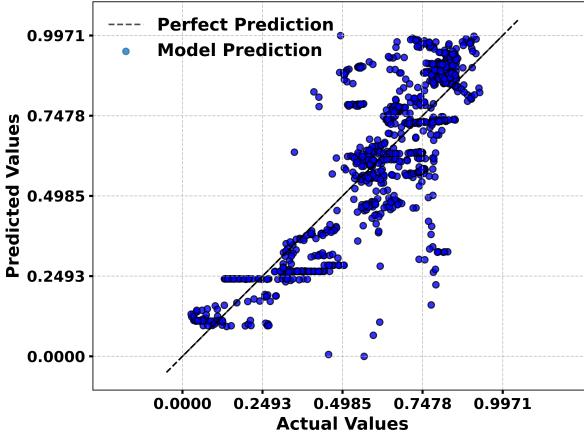


Figure 15: SVM with Polynomial Kernel. This figure presents the comparison of actual versus predicted values generated by the Support Vector Machine (SVM) utilizing a polynomial kernel, illustrating its performance and predictive accuracy.

space, the RBF kernel SVM effectively identifies optimal decision boundaries, enabling it to capture intricate patterns that simpler linear models may miss [12, 15]. This capability is crucial for modeling complex interactions among features, making the RBF kernel a powerful tool for classification and regression tasks.

Table 12 outlines the performance metrics for the RBF kernel SVM, providing insights into its predictive capabilities. The Mean Squared Error (MSE) is recorded at 0.0215, indicating a relatively low predictive error, while the Mean Absolute Error (MAE) is measured at 0.1152, reflecting the average error magnitude in the model's predictions. The R² score of 0.5772 suggests that approximately 57.72% of the variance in the target variable is explained by this model, highlighting a moderate degree of explanatory power. Additionally, the Explained Variance Score of 0.5820 underscores the model's limitations in fully capturing dataset variability, indicating room for enhancement in predictive accuracy.

The scatter plot in Figure 16 visually contrasts actual versus predicted values for the RBF Kernel SVM model, offering insight into its performance and precision. In this plot, it is observed that the points are dispersed around this line, with clusters in certain areas and noticeable deviations elsewhere, reflecting the model's limitations in capturing all patterns within the dataset.

These metrics illustrate that while the RBF kernel SVM is well equipped to manage complex data structures, its performance emphasizes the necessity for careful parameter tuning and exploration of alternative modeling techniques to improve predictive precision. Such considerations are essential for optimizing model effectiveness in real world applications.

Table 12: Performance Metrics for the SVM RBF Kernel

Sq.No.	Metric	Value
1	Mean Squared Error (MSE)	0.0215
2	Mean Absolute Error (MAE)	0.1152
3	R ² score	0.5772
4	Explained Variance Score	0.5820

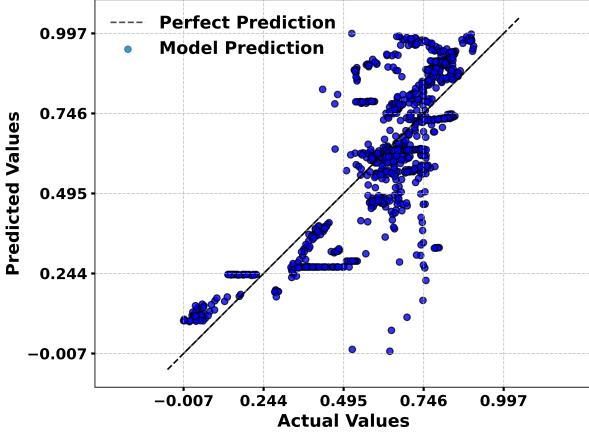


Figure 16: SVM with RBF Kernel. This figure shows the comparison of actual versus predicted values generated by the SVM using a radial basis function (RBF) kernel, illustrating its performance and predictive accuracy.

7.8. KNN

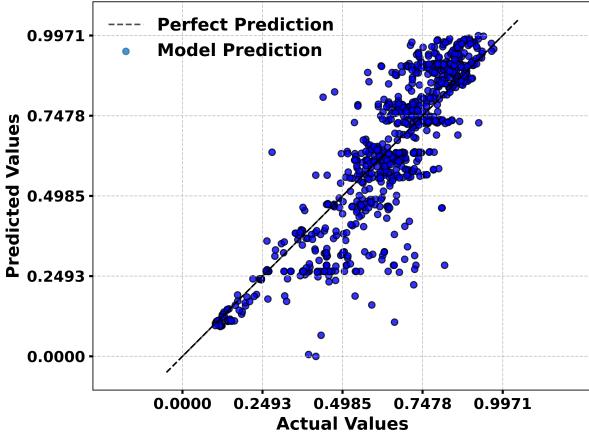


Figure 17: K-Nearest Neighbors ($k=5$). This figure shows the comparison of actual versus predicted values generated by the K-Nearest Neighbors algorithm with $k=5$, illustrating its performance and predictive accuracy.

The K-Nearest Neighbors (KNN) algorithm is a straightforward yet powerful technique widely used for classification and regression tasks. Distinct from parametric models, KNN does not assume a specific data distribution; instead, it capitalizes on the inherent structure of the dataset to make predictions [16]. The algorithm operates by identifying the nearest data points to a given query, using the majority class among these neighbors for classification or averaging their values for regression, making it particularly effective in datasets characterized by clear clusters [16].

Table 13 summarizes the performance metrics for the KNN model, shedding light on its predictive effectiveness. The Mean Squared Error (MSE) is recorded at 0.0133, indicating a low level of predictive error, while the Mean Absolute Error (MAE) stands at 0.0819, reflecting the average magnitude of errors in predictions. The R^2 score of 0.7458 suggests that approximately 74.58% of the variance in the target variable can be explained by the KNN model, demonstrating robust explanatory power. Additionally, the Explained Variance Score of 0.7495 reinforces this finding, indicating the model's effectiveness in capturing significant variability within the dataset.

The scatter plot in Figure 17 illustrates the actual versus predicted values for the K-Nearest Neighbors (KNN) model, offering a visual representation of the model's prediction accuracy. In this plot, it is observed,

Table 13: Performance Metrics for the k-Nearest Neighbors (KNN) Model with $k = 5$

Sq.No.	Metric	Value
1	Mean Squared Error (MSE)	0.0133
2	Mean Absolute Error (MAE)	0.0819
3	R^2 score	0.7458
4	Explained Variance Score	0.7495

a large concentration of points lies close to this line, especially in the middle value range, indicating the model's tendency to predict accurately within certain intervals. However, some dispersion away from the line suggests that the model struggles with certain data points, especially those further from the mean.

These metrics confirm KNN's position as a strong performer in predictive modeling. While it may not achieve the accuracy of more advanced algorithms, its simplicity and interpretability make it a valuable option for classification and regression tasks, especially in contexts where understanding the model's decision-making process is essential.

8. Ensemble Techniques for Enhanced Antenna Performance Prediction

Ensemble methods constitute a powerful suite of machine learning techniques that synthesize the predictions of multiple models to achieve enhanced accuracy and resilience [5, 9]. The core concept of ensemble learning revolves around aggregating outputs from various models, which helps reduce the likelihood of overfitting while simultaneously boosting the overall generalization capability of the predictive framework [9, 17]. This strategy proves particularly advantageous when individual models, often referred to as "weak learners," may struggle to capture the intricate complexities inherent in the data on their own [5, 17].

In this section, we will examine three prominent ensemble techniques: Decision Trees, Random Forests, and Hist Gradient Boosting. Each of these methods leverages the strengths of multiple models to enhance predictive performance, enabling them to effectively handle diverse data patterns and improve overall accuracy in various applications. By analyzing their methodologies and effectiveness, we can better understand how these ensemble techniques contribute to robust predictive modeling.

8.1. Decision tree

Table 14: Performance Metrics for the Decision Tree Model

Sq.No.	Metric	Value
1	Mean Squared Error (MSE)	0.0102
2	Mean Absolute Error (MAE)	0.0287
3	R^2 score	0.8455
4	Explained Variance Score	0.8462

A Decision Tree is a versatile non-parametric predictive model that generates predictions through a series of straightforward decision rules based on input data [18]. Its tree-like structure comprises internal nodes that represent decision criteria based on specific features, branches that illustrate the outcomes of those decisions, and leaf nodes that signify the predicted results [18]. This clear and intuitive arrangement not only enhances interpretability but also mirrors the logical decision making process, making decision trees an appealing choice for many data analysis tasks.

Table 14 presents the key performance metrics for the decision tree model, demonstrating its robust predictive capabilities. The model achieves a high R^2 score of 0.8455 and Explained Variance Score (EVS) of 0.8462, indicating a strong ability to explain the variance in the target variable. Additionally, it exhibits low Mean Absolute Error (MAE) of 0.0287 and Mean Squared Error (MSE) of 0.0102, reflecting its effectiveness in generating accurate predictions. While the Decision Tree performs admirably, it slightly trails behind the Hist Gradient Boosting model in terms of overall accuracy, suggesting opportunities for enhancement.

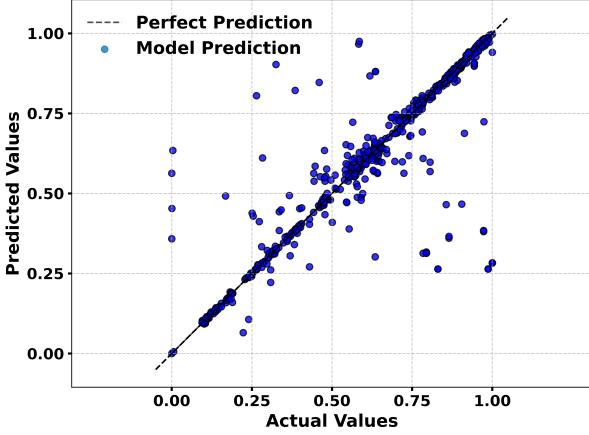


Figure 18: Decision Tree Model. This figure shows the comparison of actual versus predicted values generated by the Decision Tree model, illustrating its performance and predictive accuracy.

Figure 18 illustrates the relationship between actual and predicted values for the decision tree model. In this plot, it is observed that the majority of points align closely with the dashed line, especially near the center of the plot, signifying the model's effective performance within certain value ranges. However, a scattering of points further from the line suggests that the model struggles with more extreme values, reflecting areas where predictive accuracy declines.

These metrics reinforce the decision tree's effectiveness as a predictive tool. Although it delivers reliable predictions and maintains interpretability, its tendency to overfit in noisy or complex datasets and its relative underperformance compared to advanced models like Hist Gradient Boosting highlight the need for careful consideration in model selection and tuning to achieve optimal results.

8.2. Hist Gradient Boosting

Hist Gradient Boosting is a powerful ensemble technique that combines multiple weak learners to create a highly accurate predictive model. Its "hist" designation signifies the utilization of histogram-based strategies, which significantly enhance training efficiency, particularly with large datasets [19]. The methodology involves training decision trees sequentially, with each tree specifically designed to correct the errors of its predecessors [19]. This iterative approach enables the model to achieve high accuracy while effectively capturing complex data patterns, making it especially effective in challenging scenarios.

Table 15 outlines the performance metrics for the hist gradient boosting model, showcasing its exceptional predictive capabilities. The model achieves the highest R² score of 0.9137 and Explained Variance Score of 0.9147, demonstrating its ability to explain the majority of variance in the target variable. Furthermore, it reports the lowest Mean Absolute Error (MAE) of 0.0374 and Mean Squared Error (MSE) of 0.0051 among all evaluated models, solidifying its status as the most accurate in terms of prediction quality. These metrics highlight the model's effectiveness in handling intricate data relationships while minimizing errors.

Figure 19 displays the actual versus predicted values for the Hist Gradient Boosting model. In this plot, it is observed the data points tightly cluster around the perfect prediction line, indicating the model's high accuracy in capturing the relationship between features and the target variable. This alignment suggests that the model effectively minimizes prediction errors and consistently achieves reliable outputs, even across a range of values.

These results affirm the hist gradient boosting model as a leading choice for predictive modeling, particularly in complex datasets. Its combination of high accuracy, robustness, and efficiency underscores the importance of employing advanced techniques to achieve optimal predictive performance, setting a benchmark for future modeling efforts.

8.3. Random Forest

The Random Forest model is a highly regarded ensemble learning technique that enhances the traditional decision tree approach by constructing multiple trees during the training process [20]. This model mitigates

Table 15: Performance Metrics for the Hist Gradient Boosting

Sq.No.	Metric	Value
1	Mean Squared Error (MSE)	0.0051
2	Mean Absolute Error (MAE)	0.0374
3	R ² score	0.9137
4	Explained Variance Score	0.9147

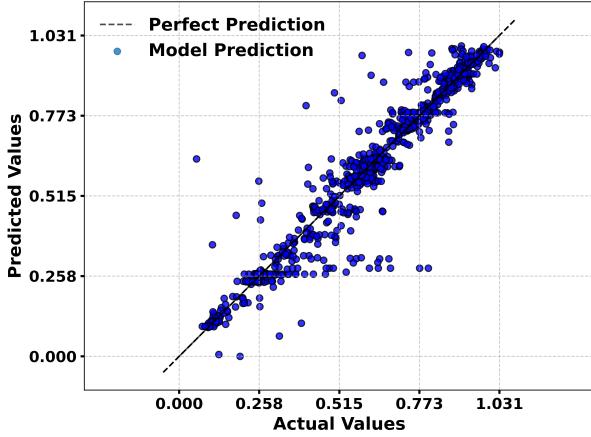


Figure 19: Hist Gradient Boosting Model. This figure shows the comparison of actual versus predicted values generated by the Hist Gradient Boosting model, illustrating its performance and predictive accuracy.

the overfitting that often plagues individual decision trees by averaging the predictions across numerous trees. Each tree is trained on a unique random subset of the data, which significantly enhances the model's generalization capabilities [20]. This ensemble methodology not only improves predictive performance but also increases robustness, making the model more resilient to noise within the dataset.

Table 16 presents the performance metrics for the random forest model, which indicate exceptional predictive capabilities. The model achieves an R² score of 0.9047, signifying that it explains approximately 90.47% of the variance in the target variable, showcasing its effectiveness in making accurate predictions. Additionally, the Mean Absolute Error (MAE) of 0.0274 and Mean Squared Error (MSE) of 0.0061 reflect minimal deviation between the predicted and actual values. The Explained Variance Score of 0.9056 further confirms the model's reliability in capturing the variability within the data, underscoring its strong performance.

Figure 20 illustrates the predicted versus actual values for the Random Forest model. In this plot, it is observed that the data points tightly cluster around the perfect prediction line indicating that the model performs well, with minimal deviation from the actual values across the dataset. This alignment suggests that the Random Forest model is effective at capturing the underlying patterns in the data, demonstrating its predictive strength.

These metrics affirm the random forest model as a premier choice for predictive tasks, particularly in complex datasets. Its impressive accuracy, coupled with low error rates, positions it as a valuable tool in ensemble modeling. The combination of high explanatory power and robustness makes the random forest model a critical component of the modeling strategies employed in this study, complementing other advanced techniques in achieving optimal predictive performance.

9. Key Insights and Implications for Comparative Model Analysis

The comparative analysis of different regression models, based on metrics such as R² score, Mean Absolute Error (MAE), Mean Squared Error (MSE), and Explained Variance Score, has provided valuable insights

Table 16: Performance Metrics for the Random Forest Model

Sq.No.	Metric	Value
1	Mean Squared Error (MSE)	0.0061
2	Mean Absolute Error (MAE)	0.0274
3	R ² score	0.9047
4	Explained Variance Score	0.9056

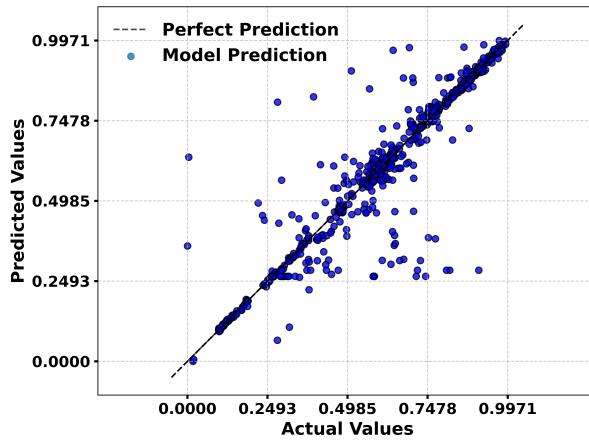


Figure 20: Random Forest Model. This figure shows the comparison of actual versus predicted values generated by the Random Forest model, illustrating its performance and predictive accuracy.

Table 17: Comparative performance of machine learning models for antenna design

Model	MSE	MAE	R ² score	Explained Variance Score
Neural Network	0.0059	0.0302	0.8925	0.8999
Dummy Model	0.0618	0.2102	-0.0011	0.0000
Linear Model	0.0296	0.1348	0.4000	0.4009
Ridge Regression	252.2769	12.2441	-8595.4128	-19.6920
Bayesian Ridge	0.0295	0.1345	0.4042	0.4051
SVM Linear Kernel	0.0293	0.1328	0.4050	0.4076
SVM Polynomial Kernel	0.0198	0.1060	0.6082	0.6119
SVM RBF Kernel	0.0215	0.1152	0.5772	0.5820
KNN (k = 5)	0.0133	0.0819	0.7458	0.7495
Decision Tree	0.0102	0.0287	0.8455	0.8462
Hist Gradient Boosting	0.0051	0.0374	0.9137	0.9147
Random Forest	0.0061	0.0274	0.9047	0.9056

into their performance on the same dataset. This section synthesizes the findings and identifies the best and worst performing models, while also discussing potential reasons for the observed outcomes.

9.1. Best Performing Models

Among the models evaluated, the Hist Gradient Boosting method emerged as the most effective, delivering outstanding results across the board. Specifically, the model achieved an R^2 score of 0.9137, which indicates that approximately 91.37% of the Variance in the target variable was explained by the model. Furthermore, it recorded a remarkably low MAE of 0.0374 and an MSE of 0.0051, reflecting the model's accuracy and precision in predictions. The high Explained Variance Score of 0.9147 further confirms the model's robustness and its ability to capture the underlying patterns in the data. These metrics suggest that Hist Gradient Boosting is exceptionally well suited for this particular application, offering both accuracy and reliability.

9.2. Moderately Performing Models

The K-Nearest Neighbors (KNN) model with $k = 5$ also showed strong performance, with an R^2 score of 0.7458, MAE of 0.0819, and MSE of 0.0133. While not as precise as the hist gradient boosting model, KNN still provides a reliable alternative, particularly in scenarios where a simpler, more interpretable model is preferred. Additionally, the polynomial SVM and linear regression models exhibited reasonable accuracy, with R^2 scores of 0.6082 and 0.4000 respectively. These models strike a good balance between being complex enough to make accurate predictions and still being efficient enough to run without heavy computational demands. This makes them a great choice when you need reliable results but also want to keep things running smoothly and quickly.

9.3. Worst Performing Models

At the other end of the spectrum, the ridge regression model performed poorly, with an R^2 score of -8595.413, a MAE of 12.2441, and an MSE of 252.277. The significantly negative R^2 score indicates that the model's predictions are worse than a simple mean prediction, suggesting a severe overfitting issue or an improper regularization parameter. Similarly, the dummy regressor showed an R^2 score close to zero, which aligns with its role as a baseline model; it offers no meaningful predictive power and serves only as a reference point for evaluating other models.

9.4. Discussion

The evaluation of various regression models presented in Table 17 reveals substantial differences in performance metrics, with certain models demonstrating remarkable predictive accuracy while others underperform. These disparities can be traced back to several key factors related to the models' underlying methodologies and their adaptability to the data.

The hist gradient boosting model excels in capturing complex, non-linear relationships within the dataset, outperforming simpler models such as linear regression and dummy regressor. Its ability to iteratively refine predictions through the sequential training of decision trees allows it to effectively minimize errors, making it particularly adept at handling large and intricate datasets. This model achieved the highest R^2 score of 0.9137, indicating that it explains a significant portion of the variance in the target variable. Additionally, its low Mean Absolute Error (MAE) of 0.0374 and Mean Squared Error (MSE) of 0.0051 highlight its precision in making predictions.

In contrast, the ridge regression model exhibits a starkly poor performance, as evidenced by its extraordinarily high MSE of 252.2769 and negative R^2 score. This suggests that the model may have been overly penalized due to an excessively high regularization strength, leading to the underfitting of the data. Such over regularization can prevent the model from capturing essential patterns and relationships, resulting in inadequate predictive capability.

The variability in performance among the SVM models, especially between the RBF and linear kernels, underscores the critical importance of selecting the appropriate kernel based on the dataset's characteristics. The RBF kernel, which is more suited for capturing non-linear patterns, outperformed the linear kernel, which assumes linear relationships that may not be present in the data.

The K-Nearest Neighbors (KNN) model showed commendable results, especially when compared to other non-linear models, due to its effectiveness in identifying local patterns. However, it requires considerable computational resources as the dataset grows, which can impact its practical applicability.

Random forest, another ensemble learning technique, also demonstrated strong performance metrics, with an R^2 score of 0.9047, MSE of 0.0061, and MAE of 0.0274. Its ability to aggregate predictions from multiple decision trees trained on random subsets of the data enhances its robustness and reduces the risk of overfitting. This model effectively captures complex relationships and interactions among features, making it suitable for diverse datasets.

Overall, the analysis highlights the necessity of carefully selecting models and fine-tuning hyperparameters to optimize performance. While hist gradient boosting emerged as the top performing model, other models such as random forest, KNN, and SVM with a polynomial kernel also achieved strong results under favorable conditions. Conversely, models like ridge regression and dummy regressor struggled, illustrating the risks of overfitting and the limitations inherent in simpler modeling approaches. Table 17 presents a comprehensive comparison of multiple machine learning algorithms based on metrics such as MSE, MAE, R^2 , and Explained Variance Score. This table offers valuable insights into the effectiveness of each model for antenna optimization. Future efforts should focus on further hyperparameter optimization and the exploration of advanced modeling techniques to enhance predictive accuracy even further.

10. Conclusion

This study presents a highly efficient dual-band slotted antenna, designed to operate at 7.7131 THz and 9.0298 THz, with notable advancements driven by machine learning techniques. The antenna features a rectangular slotted patch, fabricated on copper with a polyimide substrate, a high performance dielectric material characterized by a relative permittivity of 3.5 and a low dissipation factor of 0.006. The optimized design achieved impressive performance metrics, including a return loss of -48.974 dB, an extensive bandwidth of 848.934 GHz (spanning 7.389–8.238 THz), and a gain of 8.102 dB, while maintaining a Voltage Standing Wave Ratio (VSWR) below 2 across the impedance bandwidth. A key innovation in this research was the application of neural network models to predict critical antenna parameters such as resonant frequency, return loss, and gain facilitating a more efficient and precise design process. By conducting a comprehensive parametric sweep using CST Microwave Studio, a diverse dataset was generated, enabling an exhaustive exploration of the design space. The dataset was preprocessed rigorously, including error correction and MinMax scaling, ensuring consistency for machine learning training. A comparative analysis of machine learning models revealed that hist gradient boosting outperformed others, achieving an R^2 score of 0.9137, with a low Mean Absolute Error (MAE) of 0.0374 and Mean Squared Error (MSE) of 0.0051. The K-Nearest Neighbors (KNN) model also demonstrated solid performance, with an R^2 score of 0.7458, while polynomial SVM and linear regression provided reasonable accuracy in certain scenarios. Furthermore, the random forest model delivered robust results with an R^2 score of 0.9137, MAE of 0.0274, and MSE of 0.0061, indicating its ability to effectively capture the data's complexity. In contrast, ridge regression exhibited severe overfitting, reflected in its highly negative R^2 score, emphasizing the importance of proper hyperparameter tuning. This study underscores the transformative potential of machine learning, particularly neural networks and ensemble methods like random forest, in optimizing complex antenna systems. The research not only highlights the significant reduction in time and computational resources enabled by AI driven design approaches but also lays a strong foundation for future advancements in wireless communication systems. By demonstrating how AI driven approaches can substantially reduce the time and computational resources traditionally required for antenna design, we pave the way for more sophisticated methodologies in wireless communication systems. The insights gained from this study contribute significantly to the ongoing evolution of antenna design practices, indicating a promising future where machine learning and AI will play an integral role in engineering innovations.

References

- [1] Yildiz, C., Gultekin, S., Guney, K., & Sagiroglu, S. (2002). Neural models for the resonant frequency of electrically thin and thick circular microstrip antennas and the characteristic parameters of asymmetric coplanar waveguides backed with a conductor. *AEU-International Journal of Electronics and Communications*, 56(6), 396-406.

- [2] Sarker, N., Podder, P., Mondal, M. R. H., Shafin, S. S., & Kamruzzaman, J. (2023). Applications of Machine Learning and Deep Learning in Antenna Design, Optimization and Selection: A Review. *IEEE Access*.
- [3] Okan, T. (2021). High efficiency unslotted ultra-wideband microstrip antenna for sub-terahertz short range wireless communication systems. *Optik*, 242, 166859.
- [4] Aloui, R., Zairi, H., Mira, F., Llamas-Garro, I., & Mhatli, S. (2022). Terahertz antenna based on graphene material for breast tumor detection. *Sensing and Bio-Sensing Research*, 38, 100511.
- [5] Shi, D., Lian, C., Cui, K., Chen, Y., & Liu, X. (2022). An intelligent antenna synthesis method based on machine learning. *IEEE Transactions on Antennas and Propagation*, 70(7), 4965-4976.
- [6] Mahendran, K., Gayathri, R., & Sudarsan, H. (2021). Design of multi band triangular microstrip patch antenna with triangular split ring resonator for S band, C band and X band applications. *Microprocessors and Microsystems*, 80, 103400.
- [7] Sharma, K., & Pandey, G. P. (2021). Efficient modelling of compact microstrip antenna using machine learning. *AEU-International Journal of Electronics and Communications*, 135, 153739.
- [8] Khan, M. M., Hossain, S., Mozumdar, P., Akter, S., & Ashique, R. H. (2022). A review on machine learning and deep learning for various antenna design applications. *Heliyon*, 8(4).
- [9] Li, X., Jiang, H., Niu, M., & Wang, R. (2020). An enhanced selective ensemble deep learning method for rolling bearing fault diagnosis with beetle antennae search algorithm. *Mechanical Systems and Signal Processing*, 142, 106752.
- [10] Ghalamkari, B., & Mokhtari, N. (2022). Wide-band octagon-star fractal microstrip patch antenna for terahertz applications. *Optik*, 259, 168990.
- [11] Kalva, N., & Kumar, B. M. (2022). Feedline design for a series-fed binomial microstrip antenna array with no sidelobes. *IEEE Antennas and Wireless Propagation Letters*, 22(3), 650-654.
- [12] Gao, J., Tian, Y., & Chen, X. (2020). Antenna optimization based on co-training algorithm of Gaussian process and support vector machine. *IEEE Access*, 8, 211380-211390.
- [13] Li, Y., Dong, W., Yang, Q., Zhao, J., Liu, L., & Feng, S. (2018). An automatic impedance matching method based on the feedforward-backpropagation neural network for a WPT system. *IEEE Transactions on Industrial Electronics*, 66(5), 3963-3972.
- [14] Prabhakar, D., Karunakar, P., Rao, S. R., & Srinivas, K. (2024). Prediction of microstrip antenna dimension using optimized auto-metric graph neural network. *Intelligent Systems with Applications*, 21, 200326.
- [15] Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297.
- [16] Cui, L., Zhang, Y., Zhang, R., & Liu, Q. H. (2020). A modified efficient KNN method for antenna optimization and design. *IEEE Transactions on Antennas and Propagation*, 68(10), 6858-6866.
- [17] Jain, P., Chhabra, H., Chauhan, U., Prakash, K., Samant, P., Singh, D. K., & Islam, M. T. (2023). Machine learning techniques for predicting metamaterial microwave absorption performance: A comparison. *IEEE Access*, 11, 128774-128783.
- [18] Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, 1, 81–106.
- [19] Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5), 1189–1232.
- [20] Kulkarni, S. S., Sanghai, N. P., Baishya, C., Kumar, A., & Nayak, S. K. (2024). Random forest regression for radiation pattern prediction of planar metasurface reflector antenna. *AEU-International Journal of Electronics and Communications*, 174, 155018.