RESEARCH ARTICLE OPEN ACCESS

Navigating The 5g Terrain: A Comparative Review of Machine Learning Models for Coverage Prediction

Bairy Sai Ganesh[1], N. Naveen Kumar[2]

*(Post Graduate Student, MCA, Department of Information Technology, Jawaharlal Nehru Technological University, Hyderabad, India Email: saiganeshk777@gmail.com)

** (Associate Professor, Department of Information Technology, Jawaharlal Nehru Technological University Hyderabad, India)

Abstract:
The advancement of wireless communication technologies has positioned fifth-generation (5G) networks as a cornerstone for delivering high-speed data transmission, ultra-low latency, and uninterrupted connectivity. Unlike earlier generations, 5G is designed to accommodate a vast number of connected devices, support intelligent applications, and enable real-time information exchange. This expanded scope requires meticulous network design and optimization. One of the most crucial aspects of this process is coverage prediction, which plays a vital role in maintaining consistent service quality and optimizing network resource allocation. Accurate coverage forecasting aids service providers in the effective placement of base stations, reduces coverage gaps, and improves overall service performance.

Keywords: 5G Networks, Coverage Prediction, Machine Learning, Wireless Communication, Signal Propagation, Network Optimization, Frequency Bands, Environmental Factors, User Density, Predictive Modeling, Base Station Deployment, Model Evaluation, Computational Efficiency, Intelligent Network Planning, Telecom Infrastructure.

INTRODUCTION

In today's increasingly connected world, wireless communication is essential for nearly every aspect of modern life—from business and education to entertainment and remote operations. The emergence of fifth-generation (5g) technology has significantly raised user expectations for high-speed data access, ultra-low latency, and consistent connectivity across a variety of environments. However, delivering seamless and uniform 5g coverage remains a formidable challenge, as it is affected by several dynamic factors, including signal degradation, network congestion, physical barriers, adverse weather, and interference from nearby cell towers. These influences can result in inconsistent service quality, decreased user satisfaction, and inefficient network performance.

Conventional techniques for network coverage estimation typically involve static propagation models or labor-intensive field measurements. Although these methods provide foundational insights, they often fall short in terms of flexibility, accuracy, and scalability—particularly in rapidly changing urban or rural settings. Static models lack the capacity to reflect real-time variations in network behavior, and manual site surveys are not viable for widespread deployment due to their time and resource demands.

recent progress in machine learning (ml) has introduced smarter, data-driven solutions to the problem of coverage prediction. ml algorithms are capable of analysing vast and complex datasets that include both environmental and network-specific variables. by learning patterns from historical data, these models can make reliable predictions under new and evolving conditions, offering critical input for optimizing network planning and deployment.

this work proposes a 5g coverage prediction framework utilizing a random forest classifier to analyse multiple influential factors—such as signal strength, distance from the tower, network traffic, physical obstructions, weather conditions, time, and interference levels. the model categorizes network coverage into levels such as strong, medium, or weak.

to enhance accessibility and usability, the system is implemented as a web application, featuring a flask-based backend that hosts the prediction api, and a user-friendly frontend developed with html, css, and javascript for real-time interaction.

LITERATURE SURVEY

Previous research in wireless communication has laid a strong foundation for 5g network development. Ghosh et al. (2010) provided insights into lte-advanced technologies, emphasizing coverage challenges and the role of antenna techniques, though their work lacked focus on 5g-specific and ml-driven models. Sun et al. (2016) examined path loss in urban and indoor 5G scenarios using traditional propagation models, but their approach lacked adaptability and ML integration.

Polese et al. (2018) introduced machine learning for mmWave optimization, highlighting its potential in beamforming and handover but with limited generalization to sub-6 GHz environments. Zhang et al. (2019) focused on massive MIMO and beamforming, suggesting the inclusion of data-driven prediction methods, yet they did not present practical implementations or ML comparisons.

Simsek et al. (2016) discussed the tactile internet's latency needs, advocating for adaptive learning algorithms in coverage prediction, though empirical validation was missing. Nawaz et al. (2019) explored quantum machine learning for dense networks, offering a futuristic view but lacking real-world applicability due to hardware constraints.

Wang et al. (2020) applied ensemble learning methods like Random Forest and Gradient Boosting for signal strength prediction, demonstrating higher accuracy than traditional methods, albeit with challenges related to data dependency and retraining needs.

In contrast, the current project implements a Random Forest Classifier to predict 5G signal coverage based on real-world parameters. While it shows promising results through a web-based interface, it is presently limited by dataset scope and lacks integration with live network data for real-time adaptability

I. METHODOLOGY OF PROPOSED SYSTEM

A. *Proposed System*

The proposed system leverages supervised machine learning—specifically the Random Forest Classifier—to estimate 5G signal coverage levels. It is trained using real-world network data to classify signal strength into three categories: Strong, Moderate, and Weak, based on various extracted network and environmental features.

DATA-CENTRIC APPROACH: Relies on real-world measurements rather than theoretical models or static formulas.

COMPREHENSIVE FEATURE SET: Takes into account both network-related variables (like signal power, network congestion, and interference) and environmental elements (such as distance, obstacles, and weather conditions).

USE OF RANDOM FOREST CLASSIFIER: Selected due to its strong performance with complex, nonlinear data and its reliable accuracy.

INSTANTANEOUS PREDICTIONS: A Flask-based API enables users to receive real-time signal strength predictions based on their input.

INTERACTIVE INTERFACE: A simple and intuitive front-end built with HTML, CSS, and JavaScript allows users to input data and view results effortlessly.

ADAPTABILITY: Designed to be scalable, the model can be updated with new data to stay aligned with changing network infrastructures.

C. **System *Architecture***

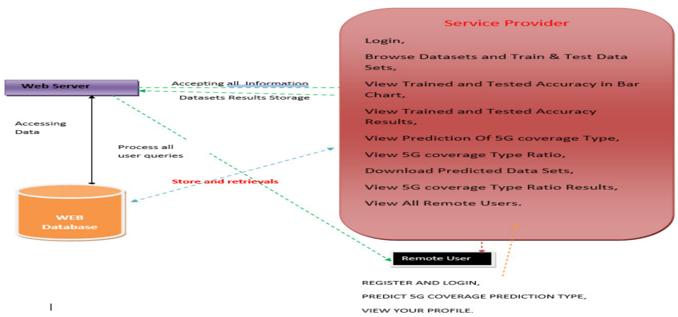


Figure-1: System Architecture of the COVERAGE PREDICTION

1. Remote User

- **Role**: An end user who can access the system remotely.
- **Main Functions**:
  - Register and Login
  - Predict 5G coverage type
  - View their own profile

➞ **Interactions:**

- Connects to the **Service Provider** component (shown with dotted arrows).

2. Service Provider

- **Role**: The main user of the platform who manages data and views analytical results.
- **Functionalities**:
  - **Login**
  - **Browse datasets** and train/test data
  - **View trained & tested model accuracy** (as chart and results)
  - **View 5G coverage predictions**
  - **Download predicted datasets**
  - **View ratio of 5G coverage types**
  - **View all remote users**

➞ **Interactions**:

- Sends data to and retrieves data from the **Web Server**
- Communicates indirectly with the **Web Database** via the server

3. Web Server

- **Role**: Central component that processes and routes all information
- **Responsibilities**:
  - Accepts all information from service providers (like datasets, results)
  - Processes all user queries
  - Manages data storage and retrieval from the **Web Database**

➞ **Key Interactions**:

- **To/from Service Provider**: Accepts inputs (datasets, commands), sends back results
- **To/from Web Database**: Retrieves and stores data as needed

4. Web Database

- **Role**: Data storage backend
- **Responsibilities**:
  - Stores datasets, results, user information, etc.
  - Responds to queries from the **Web Server**

➞ **Only interacts with Web Server**, not directly with users

🔄 **Data Flow Summary:**

1. **Remote User** registers/logs in via the **Service Provider** interface.
2. **Service Provider** accesses training/testing datasets, makes predictions, views results.
3. All requests are processed through the **Web Server**, which handles logic and data flow.
4. The **Web Database** stores and retrieves all relevant data as requested by the server.

🔁 **Key Processes:**

- **Prediction of 5G coverage type**: Triggered by user input, model is likely trained on uploaded datasets
- **Viewing ratio/results**: Analytical insights on predicted data

- **Storing/Retrieving Data**: Handled between Web Server and Database

EXPERIMENTAL ANALYSIS AND RESULTS

Results

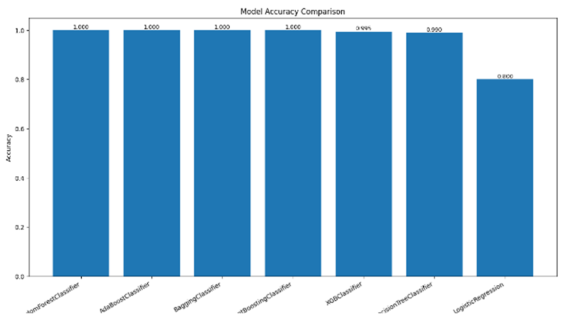
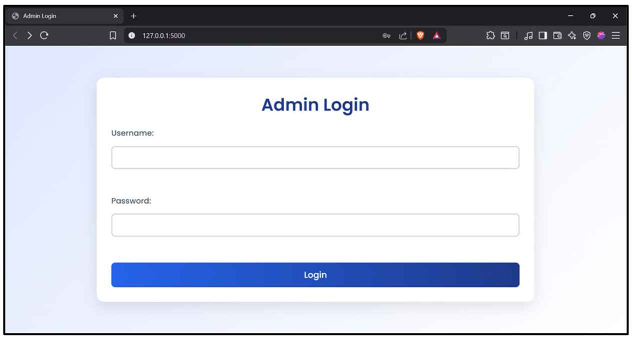


Figure-1 Accuray of different algorithm



Administrator Login Page

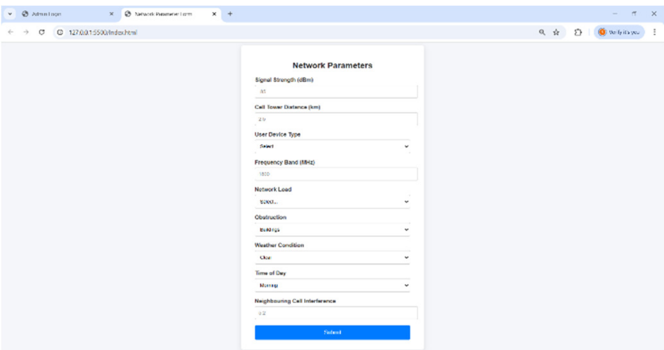


Figure-5: Form to enter the details

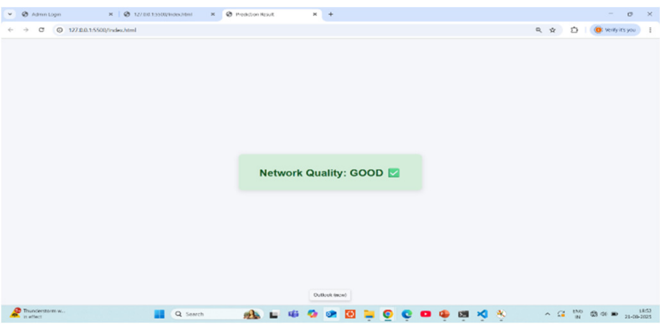


Figure-6: Result Page displaying the predictions made

CONCLUSION

This project presents an in-depth assessment of various machine learning techniques for predicting 5G network coverage, with a particular emphasis on their performance, speed, and scalability. The analysis shows that different algorithms have distinct advantages and limitations. For example, ensemble models like AdaBoost and XGBoost are highly effective with complex data due to their strong predictive capabilities. In contrast, the Random Forest Classifier (RFC) stands out for its resilience and clarity, making it ideal for extensive coverage areas. Logistic Regression (LR), while simpler in nature, offers quick and resource-efficient predictions, making it suitable for use cases where speed and simplicity are critical.

The research highlights that no single algorithm consistently outperforms the others in all situations. The optimal choice depends on factors such as the nature of the dataset, the scale of the network environment, and specific project goals. By determining the best-fitting algorithm for different real-world applications, this work demonstrates the practical benefits of using machine learning over conventional prediction models.

In conclusion, the findings underscore the value of incorporating machine learning into 5G network design. This integration can lead to smarter infrastructure deployment, more efficient resource use, and improved connectivity for users. Moreover, it addresses the scalability challenges of next-generation networks and paves the way for adaptive and intelligent wireless systems.

