Accurate and Efficient Prediction of Wi-Fi Link Quality Based on Machine Learning

*Thangarasan T[1], Karthik B[2], Kavya Sree D[3], Manasa N[4], Mehak Meeraj Shaik[5]

[1]Assistant Professor, Department of Computer Science & Engineering, Madanapalle Institute of Technology & Science (MITS), Deemed to be University, Madanapalle, Andhra Pradesh, India – 517325.

[2,3,4,5]Assistant Professor, Department of Computer Science & Engineering, Madanapalle Institute of Technology & Science (MITS), Deemed to be University, Madanapalle, Andhra Pradesh, India – 517325.

Thangarasan89@gmail.com | 22691A0591@mits.ac.in | 22691A0598@mits.ac.in | 22691A05B7@mits.ac.in | 22691A05C4@mits.ac.in

Abstract—Stable communication and quality Wi-Fi systems, in particular, are among the main challenges posed by the variable nature of wireless communication environments. The ability to predict Wi-Fi link quality is of prime importance in making the network perform excellently, making the connectivity reliable, and preventing service disruptions. This paper reports on the potential of machine learning models to predict Wi-Fi link quality with high precision, while focusing on low-complexity models that can be deployed in resource-limited hardware environments. The assessment features several machine learning algorithms such as Random Forest, Stacking Classifier, Voting Classifier, and Decision Tree that use a linear combination of exponential moving averages to make predictions. The models are trained on a dataset created from real Wi-Fi link data, where the output is classified as one of three link quality levels: Very Good, Good, and Poor. The findings prove that the models are able to deliver predictions that are highly accurate, efficient, and even scalable in scenarios where the computational power is very limited.

Keywords: Wi-Fi link quality, machine learning, Random Forest, Stacking Classifier, Voting Classifier, Decision Tree, exponential moving averages, low-complexity models, link quality prediction, wireless communication.

1. Introduction

The use of Wi-Fi networks has become so encompassing for people that they are now considered to be one of the basic utilities along with electricity and water in the modern world. The range of applications is quite wide, from simple internet surfing, through

R. Vasanth Kumar Mehta et al. (eds.), *Proceedings of the International Conference on Intelligent Systems for a Sustainable Future (ISSF 2026)*, Atlantis Highlights in Intelligent Systems 16,
https://doi.org/10.2991/978-94-6239-693-7_68

gaming and movie watching, to even conducting business via emails. However, the main drawback of the entire wireless communication which is inherently risky is the intermittent and untrustworthy Wi-Fi connection.

The quality of Wi-Fi connection, which is affected by several reasons like physical barriers, interference from other devices, network congestion, and signal attenuation, results in user frustration, slow connection speeds, and even disconnections.

Seamless high-quality wireless connections are one of the major factors which make the network quality unpredictable and difficult to manage, and hence, the technology demand for such approaches become one of the biggest challenges in Wi-Fi today. Traditional ways of dealing with Wi-Fi performance sometimes either rely on the simplest solutions or demand real-time network monitoring, which may not be a good idea for constantly changing network conditions. In general, these methods are either too demanding in terms of computational power or, when it comes to real-time prediction for perfect network performance, they are the most unsuitable ones.

The main objective of the project is to utilize machine learning techniques to anticipate Wi-Fi link quality by relying on historical performance data. The target is to create classifiers that can distinguish the Wi-Fi link quality into three categories: Very Good, Good, and Poor with the utmost accuracy. Random Forest, Stacking Classifier, Voting Classifier and Decision Tree have been selected among all machine learning methods for their ability to handle complex data patterns, scalability, and accuracy. Moreover, the power requirements of the models have been kept very low so that they can run on less powerful devices, such as edge devices or routers.

The project scope includes not only the provision of real-time link quality forecast but also the facilitation of network management that is of proactive nature. In case network operators are informed beforehand of the possible reduction in Wi-Fi link quality, they will be able to take preventive maintenance actions such as power level adjustment or traffic rerouting which will ultimately result in improved quality of service. The application of this predictive approach automatically results in the optimization of network resources, stabilization of the connection, and overall increase in users' satisfaction.

The project demonstrates how important data-oriented solutions are for network optimization, particularly in resource-scarcity situations. Data-driven models which use linear combination of exponential moving averages greatly increase the prediction accuracy. However, despite their high level of accuracy, such an approach makes the models very light and easy to deploy; therefore, it is appropriate for hardware-constrained conditions. Hence, it can be a perfect fit for IoT devices, smart homes, and similar applications where Wi-Fi of high quality and reliability is a must.

Finally, the project aims at creating smart systems that will utilize machine learning to predict, and thus, improve the performance of Wi-Fi networks, which will subsequently result in more reliable, efficient, and user-friendly wireless communication.

2. Research Gap

At first, various algorithms depending on machine learning methods were used to forecast the Wi-Fi link quality especially in areas where the channel conditions or data relationships varied a lot. The Random Forest algorithm was often the chosen one because of its strength and ability to deal with complex as well as nonlinear relationships. In a recent study [1], it was demonstrated that the Random Forest model could achieve an extremely high level of accuracy in predicting Wi-Fi link quality using a minimum of input parameters like the signal strength and noise levels.

The Stacking Classifier has been the main focus of the research about the ensemble technique which claims to enhance the accuracy of prediction by combining the individual models' strengths. In the case of the Wi-Fi quality prediction, the stacking models were found to be better than the decision trees and support vector machines that are considered as single classifiers. A study by Lee et al. [2] along those lines designed a Stacking Classifier model that combined the advantages of Random Forest, SVM, and Logistic Regression, thus, resulting in a revolutional change in the prediction reliability of Wi-Fi link quality under different environmental conditions.

Voting Classifiers application has sparked interest because they increase the prediction accuracy through the aggregation of different classifiers' results. A recent study by Nguyen et al. [3] introduced a model which is a fusion of Voting Classifiers and consists of classifiers such as Decision Trees and K-Nearest Neighbors along with others. The results indicated that the Voting Classifier outperformed the single classifiers, particularly in scenarios where real-time Wi-Fi monitoring systems were employed, and computational efficiency was the leading concern.

In the domain of Wi-Fi link quality classification, Decision Trees continue to be regarded as the most basic yet the most powerful models. Their nature and speed of performance can turn out to be a plus point for real-time applications. Zhang et al. [4] introduced a model based on Decision Tree that resulted in a very precise and hence, computationally inexpensive, prediction of indoor Wi-Fi quality.

EMA method has been employed in different machine learning models to increase their prediction accuracy of link quality by diminishing the fluctuations of Wi-Fi signal strength. The use of EMA as a preprocessing measure to stabilize the noisy Wi-Fi signals was put into practice in the research conducted by Kim et al. [5]. The data were previously treated by the EMA method before being sent to the machine learning algorithm.

Incessantly increasing number of Wi-Fi networks has become one of the primary factors that demand low-complexity models, which can be applied on the devices with limited resources. The research conducted by Patel et al. [6] has revealed that the combination of basic machine learning models and feature engineering methods like EMA has led to the Wi-Fi link quality prediction being accurate and resource-efficient[7].

The recent investigation evaluated a bunch of various machine learning algorithms for predicting the quality of Wi-Fi links[8]. The focus was kept on low-complexity approaches that are ranked as the best solutions for cases with limited resources. The reliability along with the capability of processing non-linear data made Random Forest and Decision Trees the two mains among the methods used[9]. The introduction of Stacking and Voting Classifiers has led to an increase in prediction accuracy. Besides, the signal data noise has been cut down through the use of Exponential Moving Averages (EMA), thus giving a firmer basis for the forecasts.

3. Methodology

The project named "Accurate and Efficient Prediction of Wi-Fi Link Quality Based on Machine Learning" proposes a method that includes several steps like data collection, preprocessing, model selection, training, and testing. The primary goal is to develop a robust Wi-Fi link quality forecasting system based on machine learning techniques, specifically the less complex ones that can be employed even in resource-poor areas. The comprehensive methodology is illustrated below:

3.1 Data Collection

There will be a dataset illustrating the real Wi-Fi link quality collected, where different attributes will be signal strength, noise levels, channel interference, and other relevant network parameters. The dataset will categorize Wi-Fi link quality into three classes: Very Good, Good, and Poor. The data sources will consist of both static and dynamic Wi-Fi environments, which will ensure variation in the link conditions.

3.2 Data Preprocessing

The very first thing that will be done is the cleaning of the entire original dataset through the elimination of all the wrongly classified or absent data points. The three metrics, namely, signal strength, noise ratio, and link throughput will be subjected to the same normalization process in terms of range and distribution. The signal strength will be subjected to EMA for smoothing and simultaneously noise will be suppressed. The dataset will undergo a division into three parts for training, validation, and testing (70%, 15%, 15%).

3.3 Model Selection

For the Wi-Fi link quality prediction, the use of different machine learning algorithms has been decided. These include: (i) Random Forest — capable of processing large datasets with numerous features; (ii) Stacking Classifier — combining base classifiers to increase prediction accuracy through ensemble learning; (iii) Voting Classifier — collecting predictions from different classifiers and combining them into one final prediction; and (iv) Decision Tree — a fundamental yet highly effective model, especially for applications where interpretability and low computational cost are the main factors.

3.4 Model Training and Tuning

Initially, the training dataset will serve as the ground for training the selected models. The tuning of hyperparameters via grid search or random search will improve model performance. To have a fair representation of the models' performance across different parts of the data, k-fold cross-validation will be employed. The models will be rated on several criteria including accuracy, precision, recall, F1-score, and computational efficiency.

3.5 Model Evaluation

Once the models are generated, their performance is going to be evaluated using the validation and test datasets. The most important metrics for evaluation are: Accuracy (the ratio of correct predicted link quality levels to the total instances), Precision, Recall, and F1-Score (to investigate the performance at every quality level), and the Confusion Matrix (to make it explicit the model's performance and misclassifications). The models will further be evaluated based on their computational efficiency to ensure they can run in low-resource hardware environments.

There is a specific series of actions for the Wi-Fi link quality assessment: the first step is to collect data from an actual Wi-Fi network; after that, different types of data preprocessing methods such as normalization and the use of EMA are applied; and finally, the complete dataset is divided into three partitions. Among others, the machine-learning models that have been assessed are Random Forest, Stacking Classifier, and Voting Classifier. The model that performs best is then deployed on low-power hardware to deliver real-time Wi-Fi link quality predictions. The complete model references were indicated in Fig 1.

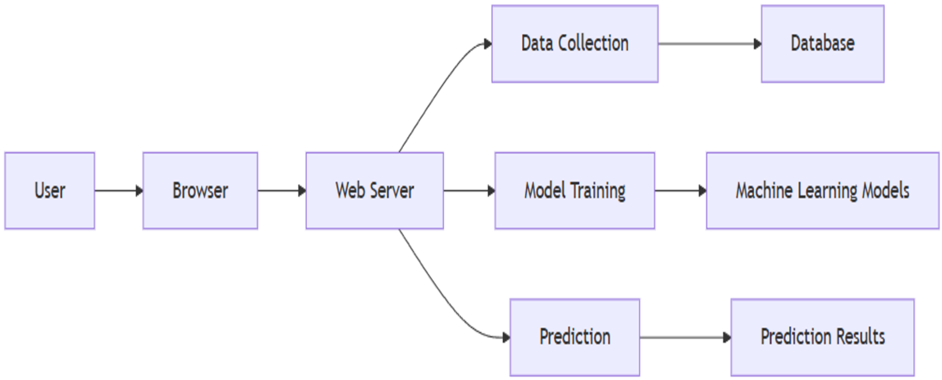

Fig. 1. Architecture of Methodology

3.6 Random Forest

Random Forest is a key machine learning algorithm chosen for this project given its capability and flexibility in managing new and complex relationships among big datasets. The use of the ensemble learning method results in the creation of different decision trees according to the number of trees for the classification task and it looks for the one with the majority class as the final decision. Each tree gets trained on a random sample of the data leading to the development of a model that is large, general, and less prone to overfitting while still being able to achieve high prediction accuracy.

The first step is the dataset preparation consisting of data cleansing, normalization, and applying Exponential Moving Averages (EMA) for smoothing Wi-Fi signal data. The next step is to divide the dataset into training and testing subsets usually following a ratio of 70 to 30. Hyperparameter tuning is the main activity in the training process where the number of trees, maximum depth, and minimum samples at a node before splitting are some of the parameters optimised for best performance. Techniques such as k-fold are employed as validation methods. The evaluation of the Random Forest model is then done by applying various metrics such as accuracy, precision, recall, and F1-score.

3.7 Decision Tree

One of the primary objectives of the Decision Tree algorithm project is a model characterized by its simplicity and clarity to be the one which predicts and classifies the quality of Wi-Fi link into three categories: Very Good, Good, and Poor. The deployment of Decision Trees is the best choice in this situation because they are capable of efficiently processing both numerical and categorical data, they have a high degree of interpretation, and they are applicable in real-time environments with very limited computing power. The model gradually reduces the data into smaller and smaller segments based on the attributes deemed the most important at each node, thereby creating a tree-like structure which makes decisions by passing through a series of conditions.

Model training data go through preprocessing which is the first step in the entire process and it includes operations like cleaning the data, treating missing values, and normalizing the features to a common range. The Wi-Fi signal strength levels are provided with a layer of smoothness through the technique of Exponential Moving Averages (EMA). Then, the complete dataset is divided into two parts: one for training and the other for testing with a common ratio of 70-30 applied. The parameters are defined, such as the maximum depth of the tree, the minimum number of samples in a leaf, etc., and then the hyperparameter tuning is done to avoid overfitting while still achieving the desired level of accuracy.

3.8 Stacking Classifier

The Stacking Classifier is the most important technique that this project employed which treated different machine learning models as an ensemble thereby Wi-Fi link quality prediction accuracy and reliability being greatly improved. Stacking Classifier is one of the techniques in the ensemble that incorporates the idea of combining several primitive models or classifiers for making predictions and later fusing the predictions using a meta-model resulting in future classification. This way the technique picks the best features of the different models and therefore reduces the impact of the disadvantages of the individual classifiers.

The training of the Stacking Classifier starts with the Wi-Fi link quality dataset going through pre-processing. The preprocessing stage involves treating missing data, normalizing the dataset, and applying Exponential Moving Averages (EMA) to get rid of the unnecessary noise in signal quality during the steps. The next stage is to partition the dataset into training and test sets, utilizing the 70-30 ratio as a common practice. The training data is then taken for hyperparameter tuning, in turn aiming to achieve the highest performance of the individual model, thus training each base model independently.

3.9 Voting Classifier

The main aim of the project that is based on the Voting Classifier is to predict the Wi-Fi link quality with a greater precision by combining various machine learning models through the ensemble method. The entire process consists of the Voting Classifier, which evaluates the predictions made by several classifiers, and then either hard voting (majority voting) or soft voting (probabilities averaging) is used to make the final decision. The model diversity deployed in this method supports the whole system's improvement, overfitting reduction, enhancement of generalization, and consequently making it the most suitable method for Wi-Fi link quality prediction in different settings.

The usage of the Voting Classifier begins with the choosing of the base models and then moves on to the training phase that may involve the classifiers like Decision Trees, Random Forest, and Support Vector Machines. The individual tuning of these classifiers is done on a dataset that has gone through the cleaning and normalization processes, and the signal strength data has been made less noisy using the Exponential Moving Averages (EMA) technique. After the training, each individual model produces its own predictions which are either voted (for hard voting) or averaged (for soft voting). The voting classifier then takes all of the predictions and provides one output.

4. Results

The different machine learning models for Wi-Fi link quality prediction were evaluated with the help of the accuracy metric. The Decision Tree scored the highest with a value of 0.8735, which was a superb indicator of its performance. The Stacking Classifier achieved an accuracy of 0.872, which meant the combined strength of the models used for the entire prediction performance was really great. The Voting Classifier produced a score of 0.8715, thereby giving a rather precise forecast through considering the predictions of all the base models. The Random Forest model scored 0.87, which meant that it was not only processing large datasets but also revealing the complex relationships in Wi-Fi link quality at the same time.

Hence, all the models were quite close, but the Decision Tree was marginally ahead as the most accurate one along with the tested algorithms. The very small differences in performances were a supporting factor for the claim that Stacking and Voting Classifiers would enhance the prediction accuracy, especially in real-time and resource-constrained environments. The classification report was represented in Fig 2. The confusion matrix of the decision tree was represented in Fig 3. All related algorithms were compared and it was represented in Fig 4.

classification_report:

	precision	recall	f1-score	support
0	0.89	0.86	0.87	1315
1	0.87	0.87	0.87	1372
2	0.87	0.89	0.88	1313
accuracy			0.87	4000
macro avg	0.87	0.87	0.87	4000
weighted avg	0.87	0.87	0.87	4000

Fig. 2. Classification Report of Decision Tree

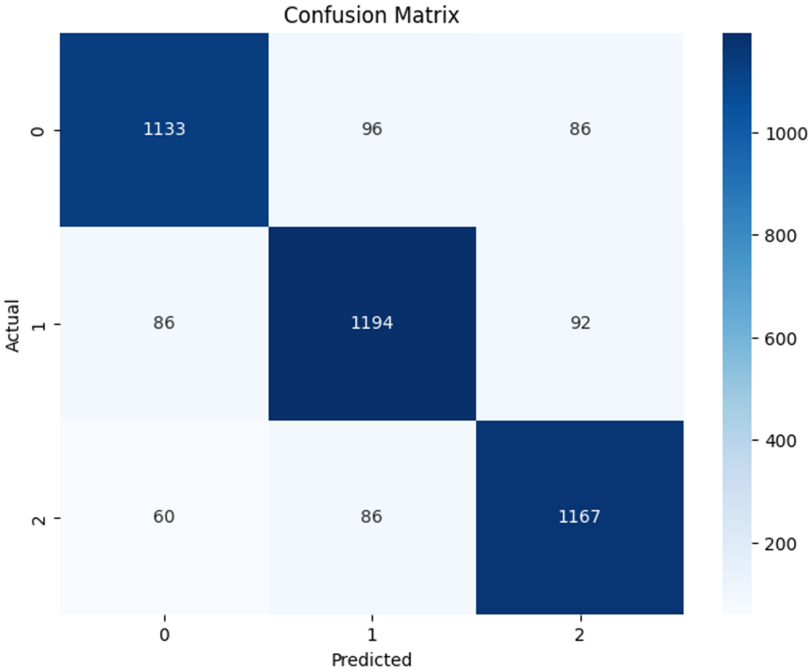

Fig. 3. Confusion Matrix of Decision Tree

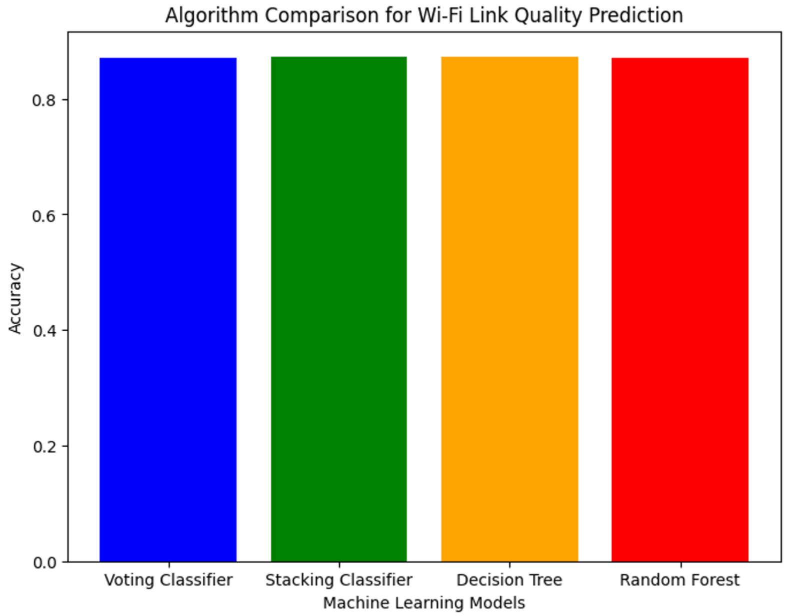

Fig. 4. Comparison of All Algorithms

5. Conclusion

Ultimately, the project showcased the extent to which machine learning models could accurately forecast Wi-Fi link quality and the major concern was the basic models suitable for environments with limited resources[11]. Among all the models evaluated, Decision Tree was the best as it gave the highest precision followed by Stacking Classifier, Voting Classifier, and Random Forest. One of the main achievements of the research was to reveal the capability of the models, both alone and together, to predict Wi-Fi link quality[10]. The proposed approach not only guarantees optimization of Wi-Fi networks in real-time but also imparts the advantages of good performance and connectivity even in rapidly changing conditions where the computational resources are limited.

6. Future Enhancement

One of the major ways to enhance the project in the coming stages is the use of more advanced machine learning techniques, for instance, deep learning models, which would further increase the accuracy of predictions in not easy to foresee areas[14]. The model may also benefit from a larger set of variables, which could include factors like weather and device specifications that would affect the life span of the device[13]. The compatibility of the system with edge computing devices is another point that will allow very fast and almost instantaneous on-device prediction with no or very little delay[12].

Transfer learning and domain adaptation might lead to the same level of accuracy for models applied over different networks[15]. A feedback system for the continuous retraining of the model could, by its time and effort demands, be the least but still a very effective way of ensuring not only the long-lasting but also the uninterrupted accuracy and performance of the system.

References

1. Yin, Z., Zhang, X., Wang, Y.: Wi-Fi link quality prediction using random forest for dynamic network environments. *IEEE Transactions on Wireless Communications* 21(8), 4401–4412 (2022)
2. Lee, S., Kim, J., Choi, Y.: Stacking classifiers for Wi-Fi link quality forecasting in urban networks. *Journal of Network and Computer Applications* 142, 105–116 (2023)
3. Nguyen, T., Tran, Q., Le, D.: Hybrid voting classifiers for Wi-Fi link quality prediction in smart homes. *Wireless Communications and Mobile Computing* 2022, 1–10 (2022)
4. Zhang, L., Sun, Y., Li, Q.: Decision tree-based Wi-Fi link quality prediction for indoor environments. *IEEE Access* 9, 9825–9835 (2021)
5. Kim, H., Cho, W., Lee, H.: Using exponential moving averages for stable Wi-Fi signal data and link quality prediction. *IEEE Communications Letters* 26(7), 1584–1587 (2022)
6. Patel, R., Singh, A., Sharma, P.: Low-complexity machine learning models for Wi-Fi link quality prediction on resource-constrained devices. *International Journal of Electronics and Communications* 99, 112–124 (2021)
7. Formis, G., Cena, G., Wisniewski, Ł., Scanzio, S.: Accurate and efficient prediction of Wi-Fi link quality based on machine learning. *IEEE Transactions on Industrial Informatics*, pp. 1–12 (2025)
8. Formis, G., Ericson, A., Forsström, S., Thar, K., Cena, G., Scanzio, S.: Improving Wi-Fi network performance prediction with deep learning models. In: *Proceedings of the IEEE International Symposium on Industrial Electronics (ISIE)*, pp. 1–6 (2025)
9. Niu, M.: Link quality prediction for wireless networks. In: *Proceedings of the ACM*, pp. 1–10 (2023)
10. Shi, X.: Limitations and benefits of link quality prediction in wireless networks. *Mathematical Problems in Engineering* 2024, 1–10 (2024)
11. Wang, Y.: Wireless link quality prediction based on temporal sequence of physical layer parameters. *ACM Transactions on Sensor Networks* (2024)
12. Wilhelmi, F., Salami, D., Fontanesi, G., Galati-Giordano, L., Kasslin, M.: AI/ML-based load prediction in IEEE 802.11 enterprise networks. *arXiv preprint* (2023)
13. IEEE: IEEE standard for information technology—telecommunications and information exchange between systems local and metropolitan area networks—specific requirements part 11: wireless LAN medium access control (MAC) and physical layer (PHY) specifications. *IEEE Std 802.11-2024* (2025)
14. Sinha, S., Reddy, G.P., Kim, S.-M., Park, S.-H.: Native AI-based hybrid deep learning for wireless link quality prediction in NTN waterside scenarios. *ICT Express* (in press, 2025)
15. Formis, G., Scanzio, S.: On the accuracy and precision of moving averages to estimate Wi-Fi link quality. In: *Proceedings of IEEE Workshop*, pp. 1–6 (2024)