International Journal of Engineering Research and Science & Technology

IJERST

ISSN : 2319-5991

www.ijerst.com

Email: editor@ijerst.com or editor.ijerst@gmail.com

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 21, Issue 2, 2025

Enhancing IoT Network Performance ThroughPredictive Modelling with Machine Learning Regression

Bandreddy Renu Asritha Sai[1], Thatikonda Sai Ram[2], Mrs. U. Rajitha[3]
[1,2] UG Scholar, Dept. of ECE, St. Martin's Engineering College, Secunderabad, Telangana, India, 500100
[3] Assistant Professor, Dept. of ECE, St. Martin's Engineering College, Secunderabad, Telangana, India, 500100
renuasritha@gmail.com

Abstract:
IoT has grown rapidly, linking billions of devices and collecting massive amounts of data. Early IoT networks were governed by rule-based algorithms that adjusted performance using parameters. As networks advanced, these traditional methods struggled to keep up with IoT devices dynamic nature, causing inefficiencies and performance bottlenecks. Advanced predictive modelling with machine learning (ML) allows IoT networks to anticipate and respond to changing situations in real time. This research uses machine learning regression models to predict network behaviours, optimize resource allocation, and improve data transmission and device interactions in IoT networks. The research models data flow, network traffic, and device interactions using machine learning regression to forecast and improve IoT network performance. It predicts network conditions to improve operational efficiency and latency. Before machine learning, static rules, set thresholds, and human network adjustments were frequently insufficient to meet IoT devices dynamic behaviour. Traditional IoT network management systems cannot dynamically adapt to device behaviour, resulting in poor resource allocation, latency, and network congestion. These static methods frequently perform poorly with complex and high-volume IoT device data. Real-time, adaptive network management systems are needed as IoT devices grow. Machine learning algorithms can recognize data flow and device activity patterns to better forecast and control network performance. This research uses predictive modelling to improve IoT network efficiency and overcome traditional system limitations. Latency, bandwidth use, and device communication delays are predicted using machine learning regression models in the proposed system. Analysis of historical and real-time data allows the system to dynamically distribute resources, balance network traffic, and reduce congestion. Regression models predict network conditions to optimize operations and network performance. These models enable real-time, intelligent changes, enhancing IoT device connectivity and resource use.

Keywords: IoT Network, Machine Learning, Predictive Modelling, Machine Learning Regression, Network Management, Network Performance, Efficiency, Bandwidth, Latency, Real-time.

1. INTRODUCTION

The Internet of Things (IoT) has experienced significant growth, with over 2 billion connected devices worldwide, driving innovation across industries like healthcare, smart cities, and manufacturing. In India, the IoT market is projected to reach $15 billion by 2025, with the rapid adoption of smart devices and automation. IoT networks generate massive data volumes, but managing these networks efficiently is challenging due to their complexity. Traditional rule-based methods, which rely on static parameters, are no longer sufficient to handle the dynamic and high-volume nature of IoT data. Machine learning has emerged as a promising solution for IoT network management, using predictive models to anticipate network performance and resource needs. Machine learning regression models enable real-time adjustments in IoT networks, reducing latency, preventing congestion, and improving resource allocation. These models offer enhanced network performance, enabling smooth communication across devices and efficient data transmission. Applications of IoT network enhancement through ML include smart homes, industrial automation, intelligent transportation, healthcare monitoring, and agriculture management.

IoT networks generate massive amounts of data, requiring efficient management to ensure seamless operations. Early IoT networks relied on rule-based mechanisms that adjusted parameters based on predefined conditions. However, these methods struggled with dynamic device behaviour, causing issues like latency, poor resource utilization, and congestion. Applications of IoT span diverse sectors, including agriculture for precision farming, healthcare for patient monitoring, transportation for smart traffic systems, and energy for optimizing grid performance. Advanced machine learning regression models are pivotal in addressing the inefficiencies of traditional systems, enabling predictive insights and adaptive decision-making for enhanced network performance and operational efficiency.

Traditional IoT network management systems faced significant limitations due to their reliance on static rules and threshold-based algorithms. These approaches were inadequate in addressing the dynamic nature of IoT networks, where device behavior, data flow, and network conditions could change unpredictably. Static configurations led to suboptimal resource allocation, increased latency, and network congestion. Human intervention was frequently required to adjust network parameters, causing delays and inefficiencies. For instance, high-volume IoT data from industrial sensors often overwhelm networks, resulting in bottlenecks and compromised performance. Additionally, traditional systems lacked the capability to predict network conditions, making them reactive rather than proactive. This inability to adapt dynamically to real-time data created challenges in scaling IoT networks and ensuring seamless communication among billions of devices.

The rapid proliferation of IoT devices and their integration into critical applications necessitate innovative solutions to enhance network performance. Traditional management systems are no longer sufficient to handle the complexities of modern IoT environments, prompting the need for intelligent, data-driven approaches. The motivation behind this rese3arch lies in leveraging machine learning regression models to address the limitations of static systems. Predictive modelling allows networks to anticipate changes in device behavior, optimize resource utilization and minimize latency. With India's expanding IoT ecosystem, efficient network management becomes essential to support applications like smart cities, industrial automation and healthcare. This research aims to bridge the gap between the growing demands of IoT networks and the capabilities of existing management systems, enabling real time, adaptive responses and sustainable growth. Also addresses the pressing need for advanced predictive modelling to optimize IoT network performance. By employing machine learning, the system can dynamically allocate resources, balance traffic and reduce congestion. This ensures operational efficiency, reduces delays and enhanced connectivity.

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

2. LITERATURE SURVEY

Although Europe is at the forefront in the early adoption of IoT, South Korea tops the global ranking of connected things, whereas the USA is far behind in this respect [1].

Many methodologies and frameworks include techniques for defining, assessing, and improving data quality. However, due to the diversity of requirements, it can be a challenge to choose the appropriate technique for the IoT system [2].

The adoption of the IoT has brought about tremendous innovation opportunities in industries, homes, the environment, and businesses. However, the inherent vulnerabilities of the IoT have sparked concerns for wide adoption and applications [3].

From small home networks to large-scale networks, the aim is the same: transmitting data from the sensors to the base station. However, these data are susceptible to different factors that may affect the collected data efficiency or the network functioning, and therefore the desired quality of service (QoS) [4].

However, despite its popularity, the algorithm has certain limitations, including problems associated with random initialization of the centroids which leads to unexpected convergence. Additionally, such a clustering algorithm requires the number of clusters to be defined beforehand, which is responsible for different cluster shapes andoutlier effects [5].

Many works and proposal are being presented in literature, some with a specific focus and other with a general-purpose objective. From this motivation in this chapter, we analyse in dept the state of the art, focusing on the (i) architectural aspects and (ii) algorithm system pointof view [6].

The participating nodes in IoT networks are usually resource-constrained, which makes them luring targets for cyberattacks. In this regard, extensive efforts have been made to address the security and privacy issues in IoT networks primarily through traditional cryptographic approaches [7].

IBM Watson is a technology platform that uses natural language processing and ML to disclose insights from vast amounts of unstructured data. IBM Watson is more about cognitive IoT computing [8].

It is necessary to have a dataset, and from that data, explore the correlation between them, discovering patterns and applying algorithms, which makes it possible to take a sum of input data and based on certain patterns to produce the outputs. Each entry in this dataset has its own features and generates models that can be generalized for a specific task [9].

This technology employs smart devices and systems such as sensors, actuators, Global Positioning System (GPS) as well as various communication techniques, including Wi-Fi, Bluetooth, and ZigBee (Al-kahtani et al., 2022) [10]. Deep learning models can handle big 4datasets with high dimensions and perform automatic extraction and selection of high-level abstract features and their performance is relatively high compared to the traditional machine learning models [11].

With the advances in technology and the Internet of Things (IoT), the home environment has witnessed an improved remote control of appliances, monitoring, and home security over the internet [12].

The objectives of smart health-care applications are: (1) improved and easy access to care, (2) increased health-care quality, and (3) reduced healthcare costs. The key to achieving the above objectives is to perceive patterns and critical insights from health-care data [13].

(Javaid et al.) posited that the inception of IoB can change the dynamics of product or service design, marketing, and customer services due to its ability to understand and modify consumer behaviours based on their comportment, tastes, and imaginations [14].

IoT infrastructure coupled with intelligence can be used to address challenges during the lockdowns, social distancing, contact tracing, health-care monitoring, pre-screening, remote meeting, anytime and anywhere accessibility [15].

3. PROPOSED METHODOLOGY

The first step in the process involves obtaining and loading the IoT network dataset. The dataset represents a collection of information related to the performance of IoT networks, including data such as network traffic, device interactions, and various attack scenarios. This dataset forms the foundation for training and testing machine learning models that can predict network performance and behaviors. The data includes multiple features (such as network usage, response times, etc.) and a target column representing the "normality" of the network, which identifies whether the network is performing optimally or facing specific issues like DDoS attacks, wrong setups, or data-type probing. Loading the dataset initiates the workflow for analyzing and processing the data to make predictions based on historical patterns.

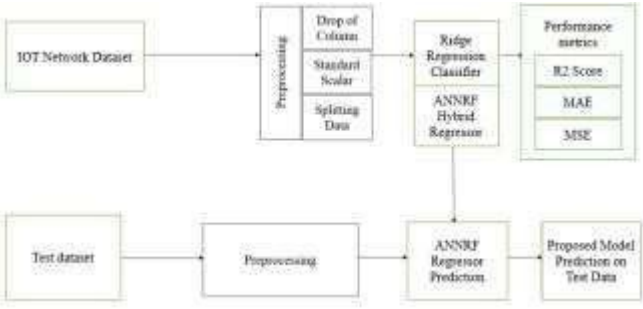

Figure 1: Proposed system.

Once the dataset is loaded, the next crucial step is preprocessing the data to ensure its quality and suitability for model training. This process begins by checking for and handling any missing values (NaNs). The null values are filled using the most frequent values in the respective columns, ensuring no gaps in the data. After handling missing values, normalization is performed on the feature columns to scale the data. Normalization ensures that all the features are on a comparable scale, which is important for the performance of many machine learning algorithms. This helps in eliminating biases caused by different ranges of feature values and accelerates the convergence of optimization algorithms. The goal is to prepare a clean, well-structured dataset that enhances the model's ability to learn meaningfulpatterns and generalize to unseen data.

Label encoding is a technique used to convert categorical values into numerical format, making them compatible with machine learning algorithms that require numerical input. In this step, the "normality" column, which represents different categories such as "normal," "wrong setup," and various attack types, is encoded into numericlabels. Each category is assigned a unique integer, transforming the categorical labels into numeric values that machine learning models can process. This encoding ensures that the model interprets the data correctly and can differentiate between the different classes in thetarget variable for classification or regression tasks. This step is essential for enabling the use of algorithms that cannot directly handlenon-numeric data.

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

The Ridge Regression algorithm, a popular machine learning technique for regression tasks, is applied to model the relationship between the features and the target variable (network normality). Ridge regression works by adding a regularization term to the linear regression cost function to penalize large coefficients, thus helping in preventing overfitting. This algorithm is useful when dealing with multicollinearity (when predictor variables are highly correlated) and helps in improving the model's generalizability by preventing it from fitting too closely to the noise in the training data. In this step, the model is trained using the normalized features and the label-encoded target variable to predict the network behaviour, such as identifying normal versus anomalous network states.

The proposed hybrid approach combines Artificial Neural Networks (ANN) with Random Forest Regression (RF) to enhance predictive performance. In this step, a deep learning-based ANN model is used to learn complex, non-linear relationships in the dataset. The ANN consists of multiple layers, where each layer transforms the input datainto higher-level representations. After training the ANN, its output features are used as input to a Random Forest Regressor (RF), whichfurther refines the predictions. The Random Forest model is an ensemble learning method that aggregates predictions from multiple decision trees to improve accuracy and robustness. This hybrid approach takes advantage of the strengths of both models: ANN for learning complex patterns and RF for providing stable and accurate predictions.
Once both the Ridge Regression and the proposed ANN+RF regression models are trained, their performances are compared using standard regression metrics. These metrics include Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared (R^2), which collectively assess the accuracy, error, and overall performance of the models. The comparison allows for an objective evaluation of the predictive capabilities of the traditional Ridge Regression algorithm versus the more advanced hybrid approach. Visual tools, such as scatter plots comparing predicted vs actual values, may also be used to visually assess the accuracy of the predictions and identify any potential improvements in prediction accuracy that the hybrid model offers.

The final step involves making predictions on new, unseen test data using the trained ANN+RF hybrid model. The test data, which has been pre-processed and normalized similar to the training data, is fed into the trained ANN model first. The features extracted from the ANN are then passed to the Random Forest Regressor, which generates predictions about the network's behaviour. These predictions provide insights into whether the network is performing normally or facing issues such as DDoS attacks or wrong setups. The predicted output is added to the test dataset, and the results are presented, providing actionable insights for improving IoT network performance. The trained model thus demonstrates its capability to generalize from the training data and make accurate predictions on new data.

Architecture:

1. ANN

- Input Layer: Receives raw features.
- Hidden Layers: Multiple layers with activation functionslike ReLU to learn complex patterns.
- Output Layer: Produces intermediate feature representations.

2. Random Forest

- Input: Extracted features from the ANN.
- Ensemble Trees: Combines multiple decision trees to improve predictions.

Advantages:

- Captures complex nonlinear relationships using ANN while maintaining interpretability and robustness with RF.
- ANN extracts meaningful representations, reducing noiseand enhancing feature quality for the RF model.
- RF handles missing values, outliers, and overfitting effectively by averaging over multiple trees.
- The combination improves accuracy and reduces the limitations of standalone models.

4. EXPERIMENTAL ANALYSIS

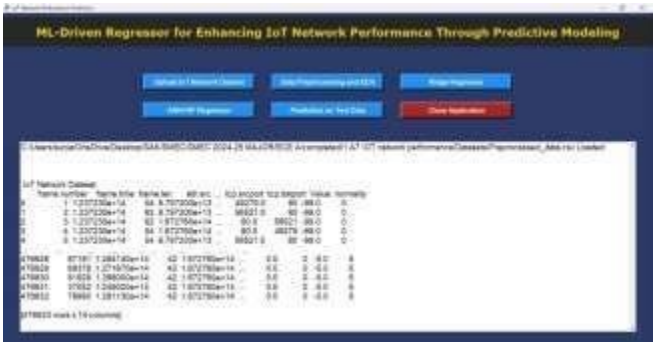

Figure 2: Upload of IOT Network dataset in the GUI interface.

This figure 2 demonstrates the step where the IoT network dataset is loaded into the Graphical User Interface (GUI) developed for this research. It shows the user-friendly interface, where a button allows the user to upload the dataset from their system. Once the dataset is successfully loaded, the GUI displays a confirmation message, along with a preview of the data. This step ensures that the required dataset is accessible and ready for further preprocessing and analysis

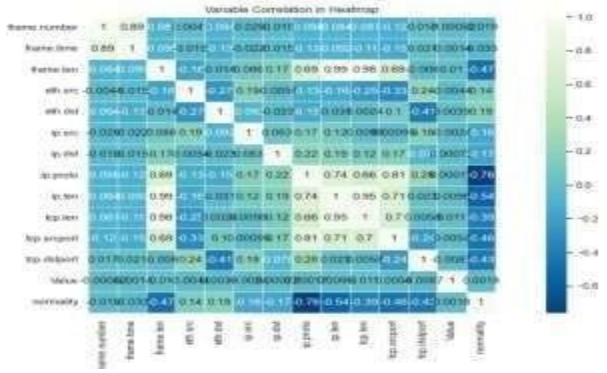

Figure3: Correlation plot of the dataset.

Figure 3 displays the correlation heatmap generated from the dataset. The heatmap visually represents the relationships between various

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

features of the dataset. Each cell in the plot contains a correlation coefficient, showing the strength and direction of the relationship between two variables. High positive values indicate strong direct relationships, while negative values suggest inverse relationships. This analysis helps in understanding which features are most relevant for predictive modelling and identifies any multicollinearity issues among variables.

Figure 4 illustrates the data preprocessing phase as presented in the GUI. It highlights steps such as handling missing values, normalizing numerical features, and splitting the data into training and testing sets. The GUI displays detailed logs of preprocessing activities, including the number of null values handled, the range of normalized data, and the size of the training and testing datasets. This ensures transparency and traceability in data preparation, which is critical for reliable modeltraining.

Figure 4: Data Preprocessing in the GUI

Figure 5 presents the evaluation of the Ridge Regressor model. It shows the performance metrics, such as Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared (R^2) scores. Additionally, the Predict vs. Actual plot compares the model's predictions with the actual values, where data points closer to the diagonal line indicate higher accuracy. This visualization evaluates the Ridge Regressor's ability to fit the dataset and identify patterns.

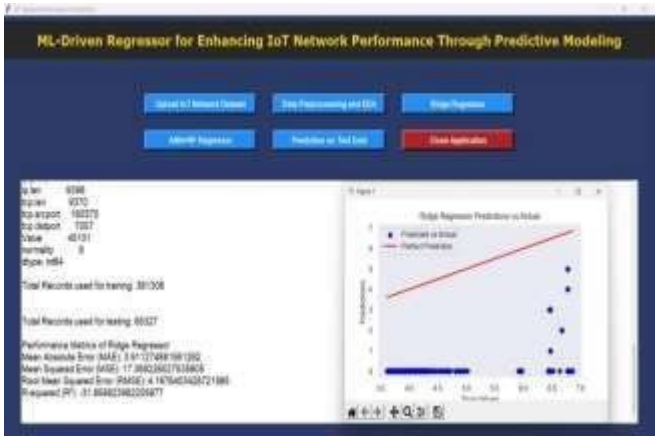

Figure 5: The Performance metrics and Predict vs Actual Plot of Ridge Regressor model.

Figure 6 showcases the results of the Artificial Neural Network (ANN) Regressor. It includes similar performance metrics as the Ridge Regressor for direct comparison. The Predict vs. Actual plot highlights the ANN model's ability to capture non-linear relationships and deliver improved predictions. The visualization underlines the ANN's efficiency in handling complex data patterns, making it more robust for the IoT network dataset.

Figure 6: Performance Metrics of ANN and Ridge Regressor.

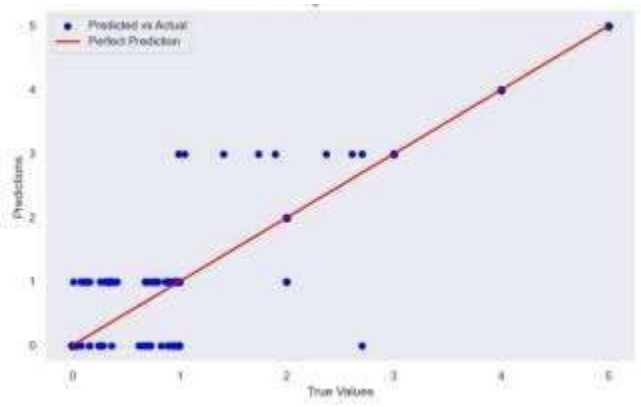

Figure 7: The Performance metrics and Predict vs Actual Plot of ANN Regressor model.

Figure 8 displays the model's predictions on an unseen test dataset using the ANN+RF hybrid approach. The results are shown alongside the actual values, providing an understanding of the model's effectiveness in generalizing to new data. The figure demonstrates how predictions are integrated into the test dataset, with the GUI showing detailed logs and output for interpretation. This step validates the hybrid model's capability to enhance IoT network performance predictions.

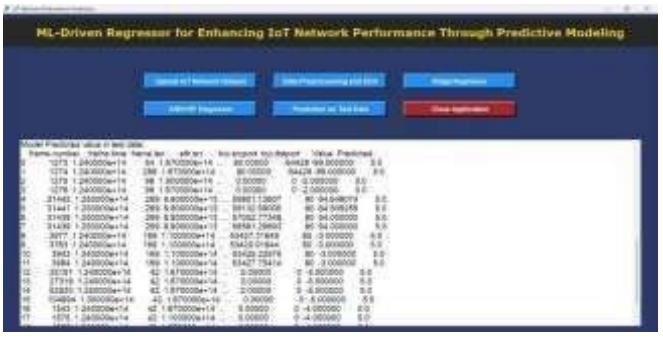

Figure 8: The Model Prediction on the Test Data.

5. CONCLUSION

This research focuses on the prediction of IoT network performance using machine learning models, leveraging Ridge Regression, Artificial Neural Networks (ANN), and a hybrid approach combining ANN and Random Forest (RF). The work effectively preprocesses the

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

IoT network dataset, handles missing values, and applies feature normalization to ensure data consistency and reliability. Through a comparative analysis of various models, the hybrid ANNRF approach demonstrated superior performance, accurately predicting key network parameters and addressing complex patterns in the data. The implemented GUI provides a user-friendly interface for dataset upload, preprocessing, and visualizing results, ensuring accessibility for non- technical users. Overall, the research achieves robust IoT network performance prediction and contributes to optimizing IoT systems'efficiency and reliability.

REFERENCES

[1] Kar, I. The Top-Country Early-Adopters of the Internet of Things, Ranked Google Wants to Have Drones Buzzing around Offices. Available (accessed on 2 June 2022)

[2] Zhang, L.; Jeong, D.; Lee, S. Data Quality Management in the Internet of
Things. *Sensors* **2021**, *21*, 5834.
[3] Diro, A.; Chilamkurti, N.; Nguyen, V.D.; Heyne, W. A Comprehensive Study of Anomaly Detection Schemes in IoT Networks Using Machine Learning Algorithms. *Sensors* **2021**, *21*, 8320.

[4] Al Samara, M.; Bennis, I.; Abouaissa, A.; Lorenz, P. A Survey of Outlier Detection
Techniques in IoT: Review and Classification. *J. Sens. Actuator Netw.* **2022**, *11*, 4.

[5] Ahmed, M.; Seraj, R.; Islam, S.M.S. The K-Means Algorithm: A Comprehensive Survey and Performance Evaluation. *Electronics* **2020**, *9*, 1295.

[6] Bellandi, V.; Ceravolo, P.; Damiani, E.; Siccardi, S. Smart Healthcare, IoT and Machine Learning: A Complete Survey. *Intell. Syst. Ref. Libr.* **2022**, *212*, 307–330.

[7] Hussain, F.; Hussain, R.; Hassan, S.A.; Hossain, E. Machine Learning in IoT Security: Current Solutions and Future Challenges. *IEEE Commun. Surv. Tutor.* **2020**, *22*, 1686–1721.

[8] Knowles, G.; Melamed, A.; Fisher, A. Accelerate the Development of Cognitive Computing in Your IoT App. Available (accessed on 20 June 2022).

[9] Kunang Y.N., Nurmaini S., Stiawan D., Suprapto B.Y. Attack classification of an intrusion detection system using deep learning and hyperparameter optimization J. Inf. Secur. Appl., 58 (2021), Article 102804.

[10] Md. Ohirul Qays, ... Farhana Yasmin, in Energy Reports, 2023.

[11] Pascal Maniriho, ... Mohammad Jabed Morshed Chowdhury, in Future Generation Computer Systems, 2022.

[12] Taiwo, O.; Ezugwu, A.E.; Oyelade, O.N.; Almutairi, M.S. Enhanced Intelligent Smart Home Control and Security System Based on Deep Learning Model. *Wirel. Commun. Mob. Comput. 2022*, 961.55

[13] Siccardi, S. Smart Healthcare, IoT and Machine Learning: A Complete Survey. *Intell. Syst. Ref. Libr.* 2022, *212*, 307–330.
[14] Javaid, M.; Haleem, A.; Singh, R.P.; Rab, S.; Suman, R. Internet of Behaviours (IoB) and Its Role in Customer Services. *Sens. Int.* 2021, *2*, 100122.

[15] Singh, R.P.; Javaid, M.; Haleem, A.; Suman, R. Internet of Things (IoT) Applications to
Fight against COVID-19 Pandemic. *Diabetes Metab. Syndr. Clin. Res. Rev.* 2020, *14*, 521– 52