

Hierarchical Resource Intelligence for Multi-Tier IoT Allocation Classification and Resource Ratio Prediction

Mohammed Aliuddin Hyder¹, K. Srilatha^{2*}

¹PG Student, ²Assistant Professor, ^{1,2}Department of Computer Science and Engineering
^{1,2}Vaagdevi Engineering College, Warangal, 506005, Telangana, India.

¹mohammedaliuddinhyder04@gmail.com

*Correspondence: K. Srilatha

ABSTRACT

With the proliferation of IoT devices, there has been exponential growth in data generation, placing substantial demands on both cloud computing (CC) and internet infrastructure. CC, renowned for its scalability and virtual resource provisioning, plays a vital role in supporting modern e-commerce and IoT applications. However, the dynamic nature of IoT environments introduces significant challenges in efficient resource allocation, service-level agreement (SLA) compliance, and continuous performance monitoring. This research presents an intelligent framework for optimizing resource allocation in multi-tier IoT environments by simultaneously performing Classification and Regression Tasks (CART). The framework analyzes device- and network-level parameters, including CPU usage, memory usage, network latency, jitter, task execution time, and sensor data, to determine optimal task allocation and predict the percentage of resource utilization. Initially, the dataset undergoes preprocessing through feature extraction, label encoding, normalization, and train-test splitting to prepare high-quality inputs for model development. Several baseline models, including Decision Tree (DT) based on CART, Random Forest (RF), Gradient Boosting (GB), and Deep Q Network (DQN), are implemented for comparative analysis. Furthermore, the proposed Hybrid BiLSTM Ensemble Learning (HBEL) model integrates a Bidirectional Long Short-Term Memory (BiLSTM) network with a probabilistic Tao Tree (PTT) as an ensemble to improve prediction robustness, stability, and accuracy through prediction fusion. The

developed models are evaluated using accuracy, precision, recall, and F1-score for CA, and MAE, MSE, RMSE, and R² score for RT. Experimental results demonstrate that HBEL consistently outperforms the baseline models, providing an effective and reliable solution for intelligent IoT resource management.

Key words: Internet of Things, Cloud Computing, Edge Computing, Resource Optimization, Hybrid Deep Learning, Classification and Regression Trees

1. INTRODUCTION

The rapid advancement of networking and communication technologies has significantly transformed the way computing resources are utilized and managed across distributed environments [1]. Early computing infrastructures primarily relied on client-server architectures, where multiple client devices communicated with centralized servers responsible for data processing, storage, and decision-making. In this architecture, end devices collected data and transmitted it through communication networks to remote servers, where computational tasks were executed. Although this approach improved accessibility, centralized management, and service availability, it also introduced several challenges, including increased network congestion [2], higher communication latency, bandwidth limitations, and strong dependence on continuous network connectivity. These limitations became particularly critical in time-sensitive applications where delays in processing could adversely affect system responsiveness and service quality. The evolution of interconnected computing systems

can be traced to traditional centralized architectures in which data generated by multiple devices or terminals was transmitted to a central processing unit for analysis and decision-making [3]. In such environments, computing resources were concentrated within data centers or centralized servers, while end devices primarily functioned as data acquisition units with limited computational capabilities. While this architecture was effective for small-scale deployments [4], the exponential growth of connected devices and data volumes exposed significant scalability and performance limitations.

In Internet of Things (IoT) ecosystems and Wireless Sensor Networks (WSNs),

conventional systems largely depended on periodic transmission of sensor data to centralized gateways or cloud servers. Resource management and task allocation were commonly performed using predefined rules, static configurations, or threshold-based mechanisms. These approaches did not adequately consider dynamic changes in network conditions, workload fluctuations, device status, or resource availability. Consequently, inefficient utilization of computational resources, excessive energy consumption, increased communication overhead, and suboptimal system performance became common challenges.

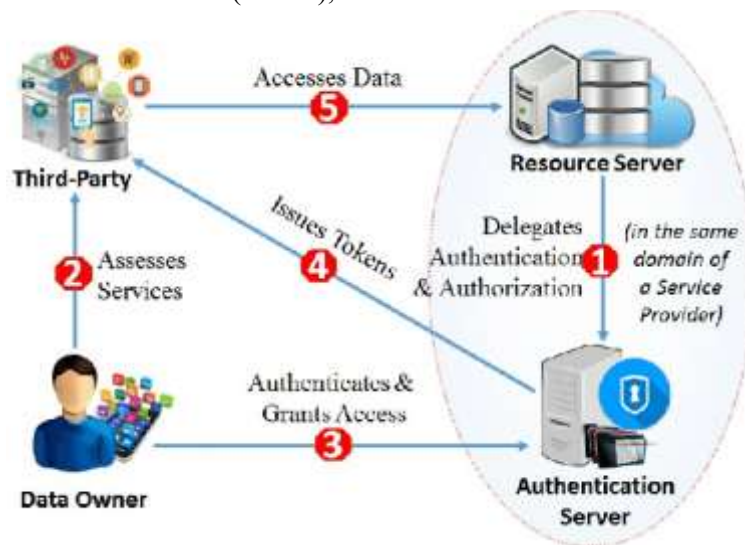


Figure 1: Traditional OAuth-Based Authorization and Access Control Framework for Resource Sharing

Figure 1 illustrates the traditional OAuth-based authorization architecture used for secure resource sharing between a data owner, third-party application, authentication server, and resource server. Initially, the resource server delegates authentication and authorization responsibilities to the authentication server. The data owner evaluates and selects a third-party application or service that requires access to protected resources. The data owner then authenticates with the authentication server, which verifies the credentials and grants access permissions. After successful authentication, the authentication server issues an authorization token to the third-party application. Using this

token, the third-party application requests and accesses protected data stored on the resource server. This framework enables controlled and secure access to resources without exposing user credentials directly to third-party services. However, the architecture relies heavily on centralized authentication and authorization mechanisms, which can introduce challenges related to scalability, single points of failure, trust management, and secure handling of sensitive data in large-scale distributed environments.

2. LITERATURE SURVEY

Alam et al. [5] proposed a systematic review framework for AI-driven resource allocation in

cloud computing by analyzing 63 high-quality studies selected through the PRISMA methodology. The methodology categorized resource allocation approaches based on algorithmic techniques, deployment environments, optimization objectives, and evaluation strategies. Quantitative analysis was conducted to compare machine learning, deep learning, reinforcement learning, and hybrid methods for cloud resource management. The study examined performance metrics such as latency reduction, cost optimization, and energy efficiency. Furthermore, sustainability aspects and evaluation practices were assessed to identify research gaps in AI-enabled cloud infrastructures. Most studies relied on simulation environments, with limited real-world deployment validation and insufficient focus on sustainability metrics. Saurabh et al. [6] developed an adaptive bandwidth scheduling algorithm for IoT devices using machine learning techniques. The methodology collected network traffic information and utilized learning-based prediction models to estimate future bandwidth requirements. Dynamic scheduling mechanisms allocated communication resources according to device demand and network conditions. Traffic prioritization strategies were incorporated to reduce congestion and improve throughput. Experimental analysis demonstrated enhanced bandwidth utilization and service quality in IoT environments. The algorithm requires continuous retraining to maintain accuracy under rapidly changing traffic patterns. Wang et al. [7] proposed an optimal federated learning-based resource allocation scheme for NOMA-enabled Industrial Internet of Things systems. The methodology integrated federated learning with non-orthogonal multiple access technology to optimize communication and computing resources. Distributed model training was performed without sharing raw data among devices. Resource allocation decisions considered transmission power, channel conditions, and learning efficiency. The framework improved network performance while preserving data privacy. Federated

learning introduces communication overhead during model aggregation and synchronization. Akashwar et al. [8] introduced an IoT-based data acquisition and machine learning forecasting framework for intelligent solar microgrid control. The methodology utilized IoT sensors to collect environmental and energy generation data in real time. Machine learning models predicted future power generation and consumption patterns. Forecasting outcomes were incorporated into microgrid control strategies for efficient energy management. The system enhanced renewable energy utilization and operational reliability. Prediction accuracy can be significantly affected by sudden environmental changes and weather anomalies. Amodu et al. [9] developed a systematic mapping study on machine learning applications in non-terrestrial network-assisted IoT systems. The methodology reviewed enabling technologies, communication architectures, and intelligent resource management approaches. Various machine learning models were analyzed for routing, spectrum allocation, and mobility management. The study categorized existing solutions based on performance objectives and deployment scenarios. Future research opportunities for integrating AI into non-terrestrial IoT infrastructures were identified. The work primarily provides a survey perspective without practical implementation or performance validation. Solomon [10] proposed an artificial intelligence-based framework for intelligent resource allocation in cloud and edge computing environments. The methodology employed machine learning algorithms to analyze workload characteristics and resource demand patterns. Predictive models enabled proactive allocation of computing resources. Edge-cloud coordination mechanisms improved latency and utilization efficiency. The framework supported dynamic adaptation to fluctuating application requirements. Resource allocation decisions depend heavily on the accuracy of workload prediction models.

Tang et al. [11] developed an IoT and machine learning-based intelligent logistics supply chain framework. The methodology integrated IoT devices for real-time monitoring of goods and transportation processes. Machine learning algorithms analyzed logistics data to optimize routing and resource allocation. Predictive analytics improved inventory management and delivery scheduling. The system enhanced operational efficiency throughout the supply chain lifecycle. Large-scale deployment requires extensive sensor infrastructure and continuous data collection. Sudha et al. [12] introduced a review of machine learning-based precision agriculture techniques for IoT-enabled crop monitoring. The methodology examined sensing technologies, data acquisition methods, and predictive analytics models used in smart farming. Machine learning algorithms were evaluated for irrigation scheduling, crop health assessment, and yield prediction. IoT devices facilitated real-time environmental monitoring. The review highlighted the role of intelligent decision support systems in modern agriculture. Most solutions face challenges related to sensor reliability and heterogeneous agricultural environments. Lin [13] proposed an IoT-assisted machine learning framework for optimizing educational resource allocation in ideological and political education. The methodology collected educational usage data through IoT-enabled systems and analyzed resource demand patterns. Machine learning models predicted learning requirements and utilization trends. Optimization strategies dynamically distributed educational resources among users. The framework improved accessibility and allocation efficiency in educational environments. The model may not generalize effectively across institutions with different educational structures. Alam et al. [14] developed a combined Q-learning methodology for dynamic resource alignment in Industrial IoT networks. The methodology employed reinforcement learning to continuously adapt resource allocation decisions according to network conditions. Q-

learning agents learned optimal policies through interactions with the environment. Resource alignment considered communication efficiency, latency, and energy utilization. The framework enhanced adaptability in dynamic industrial applications. Convergence time can become excessively long in large-scale Industrial IoT deployments. Khan et al. [15] introduced a collection of optimization techniques supporting Green IoT development. The methodology examined energy-aware communication protocols, intelligent scheduling strategies, and resource-efficient architectures. AI-based optimization mechanisms were employed to minimize energy consumption while maintaining network performance. Sustainability metrics were incorporated into resource allocation decisions. The framework contributed to environmentally responsible IoT deployments. Balancing energy efficiency and quality of service remains a challenging optimization problem.

3. PROPOSED SYSTEM

This research architecture integrates a complete pipeline for intelligent task classification and resource allocation prediction through a Flask-based application. Initially, the application loads the dataset and performs comprehensive preprocessing, including label encoding, feature scaling, and feature selection to prepare high-quality input features. As illustrated in Figure 2, the processed data is simultaneously utilized for classification and regression tasks, where multiple baseline models including DT, RF, GB, DQN, including the proposed HBEL model, are trained and evaluated. The classification module predicts the target class, while the regression module estimates the resource allocation percentage. The trained models are archived for future inference and support both single and batch prediction modes. Finally, the system presents the predicted outcomes and performance metrics through an interactive dashboard, enabling effective monitoring and decision-making.

Step 1: Application Initialization and Data Loading

The system begins by launching the Flask application, which acts as the interface for model training and prediction. The dataset is then imported from the local storage and validated to ensure consistency before processing. This stage establishes the foundation for the subsequent preprocessing and learning operations.

Step 2: Data Preprocessing and Feature Engineering

The loaded dataset undergoes preprocessing to improve data quality and model compatibility. Categorical attributes are transformed using label encoding, while numerical features are standardized through feature scaling. Relevant attributes are then selected to reduce

redundancy and improve the predictive capability of the learning models.

Step 3: Model Training for Classification and Regression

The processed dataset is simultaneously supplied to the classification and regression pipelines. Multiple machine learning algorithms, including Decision Tree, Random Forest, Gradient Boosting, DQN, and the proposed Hybrid BiLSTM-Ensemble model, are trained to perform target classification and resource allocation prediction. This parallel learning approach enables comprehensive performance comparison across different models

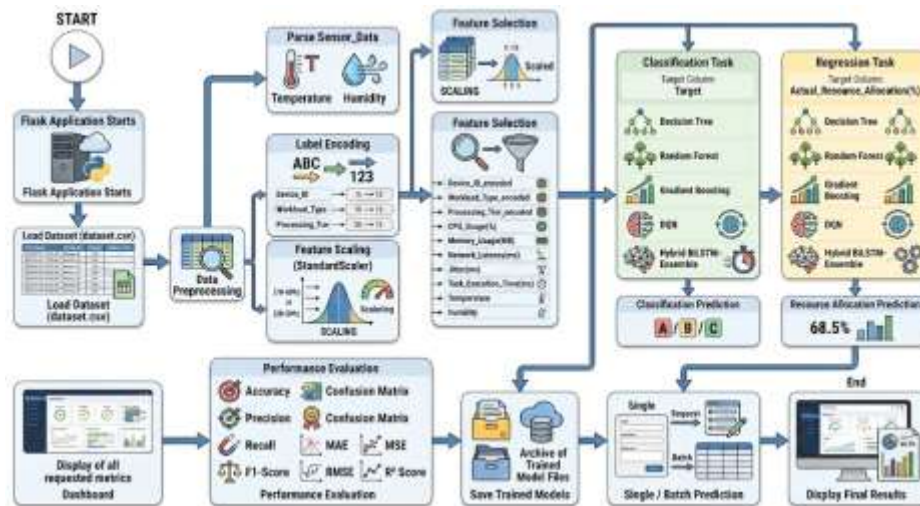


Figure 2: Proposed system architecture

Step 4: Performance Evaluation and Model Archiving

Each trained model is evaluated using appropriate performance metrics such as Accuracy, Precision, Recall, F1-Score, Confusion Matrix, MAE, MSE, RMSE, and R² Score. The evaluation results are compared to identify the most effective model for each task. The optimized models are then stored in the model repository for future deployment.

Step 5: Prediction and Result Visualization

The archived models support both single-instance and batch prediction modes for practical deployment scenarios. Based on user inputs, the system generates classification and

resource allocation predictions efficiently. The final outputs, along with the evaluation metrics, are displayed through a dashboard that provides a comprehensive view of the model performance and prediction results.

4.3.5 HBEL-CART

The HBEL model is a hybrid prediction framework as shown in Figure 3 developed to perform both classification and regression for intelligent resource allocation in multi-tier IoT environments. The model combines the sequential feature learning capability of BiLSTM with an PTT ensemble decision model to improve prediction robustness and generalization. The extracted input features are

simultaneously processed through parallel learning branches, where BiLSTM captures complex feature dependencies while the ensemble model provides complementary decision learning. The outputs from both branches are integrated using a prediction fusion mechanism to generate the final classification and regression results. This collaborative learning strategy minimizes prediction errors, improves stability, and enhances the overall resource allocation performance compared with individual machine learning and deep learning models.

Step 1: Feature Input

The HBEL model receives the normalized feature vector generated during data preprocessing. The input consists of encoded categorical attributes together with numerical

system parameters such as CPU usage, memory usage, network latency, jitter, task execution time, temperature, and humidity. These features collectively represent the operational state of IoT devices and serve as the foundation for subsequent learning.

Step 2: Parallel Feature Learning

The processed feature vector is simultaneously forwarded to the BiLSTM learning branch and the ensemble learning branch. The BiLSTM learns nonlinear relationships among the input features, whereas the ensemble branch independently captures complementary decision patterns. Parallel feature learning enables the model to exploit diverse representations without introducing redundant computations.

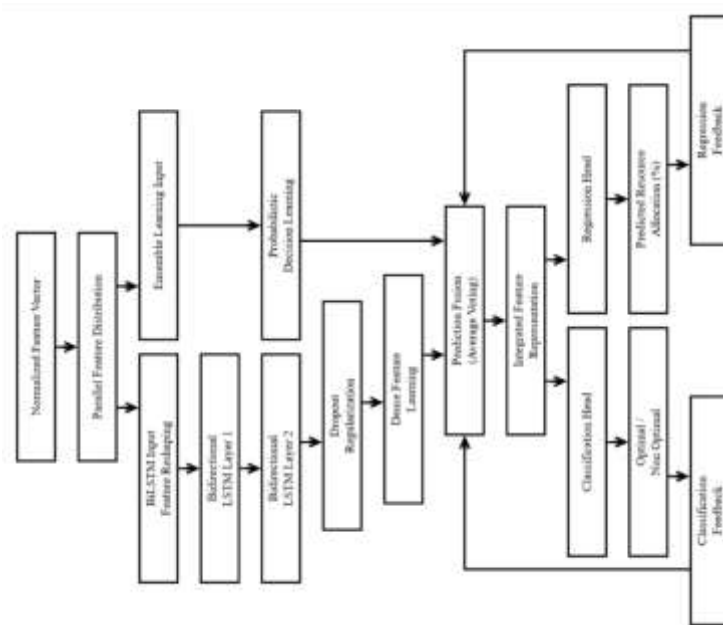


Figure 3: Operational flow of HBEL-CART

Step 3: BiLSTM Feature Extraction

Within the BiLSTM branch, the feature sequence passes through stacked bidirectional LSTM layers that analyze information in both forward and backward directions. This bidirectional learning captures hidden dependencies among resource utilization parameters and generates high-level feature representations that improve predictive capability.

Step 4: Ensemble Decision Learning

The ensemble branch receives the same normalized feature vector and performs probabilistic decision learning using multiple decision structures. This branch focuses on generating reliable predictions by reducing variance and improving decision consistency, thereby complementing the deep feature representations produced by the BiLSTM branch.

Step 5: Prediction Fusion

The outputs generated by the BiLSTM and ensemble branches are combined through a prediction fusion strategy. The fusion mechanism aggregates the predictions from both branches to produce stable and robust outputs while minimizing individual model biases and improving overall prediction reliability.

Step 6: Classification Task

The fused prediction is utilized for the classification task to determine whether the task allocation is optimal or non-optimal. The classification branch produces the final decision after integrating information learned from both the deep learning and ensemble learning components.

Step 7: Regression Task

The fused prediction is also employed for regression to estimate the percentage of resource allocation required by the IoT task. The regression output provides a continuous prediction that assists in intelligent resource

scheduling and efficient workload distribution across computing tiers.

4. RESULTS AND DISCUSSION

Figure 4 illustrates the confusion matrix of the proposed HBEL model for the offload training target, demonstrating its classification performance in distinguishing Optimal and Non-Optimal resource allocation instances. The matrix indicates that the model correctly classified 623 Non-Optimal and 556 Optimal samples, while only 21 Optimal instances were incorrectly identified as non-optimal, with no false positive predictions observed for the Non-Optimal class. These outcomes reflect the model's strong discriminative capability and its ability to minimize classification errors across both categories. The limited number of misclassifications highlights the effectiveness of the hybrid learning strategy in capturing complex feature relationships and producing reliable prediction outcomes



Figure 4: Confusion matrix using HBEL-CART model for offload training target

Figure 5 depicts the scatter plot comparing the actual and predicted resource allocation values generated by the proposed HBEL model. The distribution of data points demonstrates a strong positive correlation, with most predictions closely following the perfect prediction reference line, indicating high agreement between the observed and estimated values. Although minor deviations are visible at

certain allocation levels, the overall dispersion remains limited, reflecting the model's ability to estimate resource allocation with high precision. The consistent alignment of predictions across the complete range of values highlights the robustness of the hybrid learning framework in capturing complex relationships within the dataset

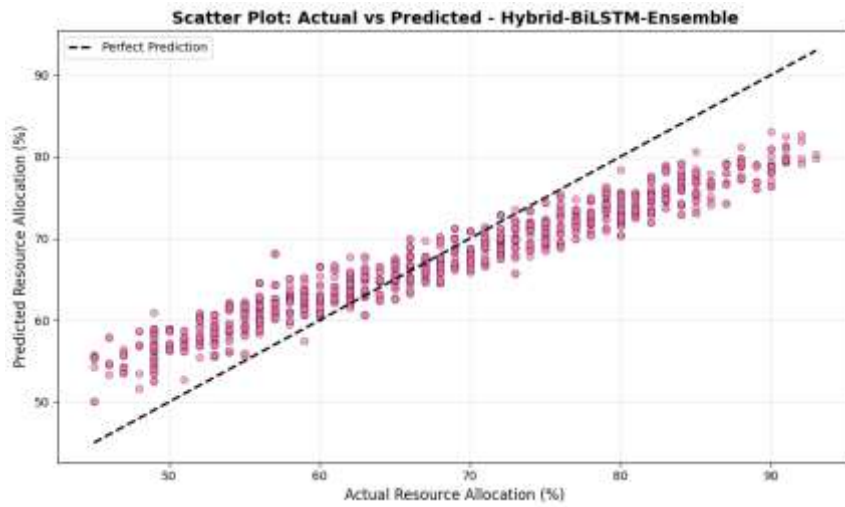


Figure 5: Illustration of scatter plot using HBEL

Figure 6 shows that the system has classified the given input as Non-Optimal, indicating that the current resource allocation or task placement is not efficient under the provided conditions. Additionally, the model predicts a resource allocation of 71.12%, suggesting the

estimated percentage of resources required for the task. These predictions are generated using the HBEL model, which combines deep learning and ensemble learning techniques, ensuring high accuracy and reliability in both classification and regression outputs.

Figure 6: prediction obtained using HBEL CART

Table 1: Performance Comparison of Classification models for offloading_train class

| Model | Accuracy | Precision | Recall | F1-Score |
|----------------------|---------------|---------------|---------------|---------------|
| Existing DT | 54.75% | 54.74% | 54.75% | 52.97% |
| Existing RF | 55.83% | 55.78% | 55.83% | 55.79% |
| Existing GB | 83.17% | 83.34% | 83.17% | 83.17% |
| Existing DQN | 96.83% | 96.83% | 96.83% | 96.83% |
| Proposed HBEL | 98.25% | 98.31% | 98.25% | 98.25% |

Table 1 presents the comparative performance analysis of the existing and proposed models using the evaluation metrics of accuracy, precision, recall, and F1-score. The existing DT achieved an accuracy of 54.75%, followed by RF-CART with 55.83%, GB-CART with 83.17%, and DQN-CART with 96.83%, demonstrating a progressive improvement in classification performance across the baseline models. The proposed HBEL CART model attained the highest accuracy of 98.25%, outperforming all the existing approaches.

Table 2: Performance Comparison of regression models for resource_allocation

| Model | MAE | MSE | RMSE | R ² -Score |
|---------------|--------|--------|-------|-----------------------|
| Existing DT | 10.044 | 138.58 | 11.77 | 0.035 |
| Existing RF | 10.040 | 137.76 | 11.73 | 0.040 |
| Existing GB | 8.412 | 98.986 | 9.94 | 0.310 |
| Existing DQN | 9.9701 | 138.21 | 11.75 | 0.030 |
| Proposed HBEL | 4.854 | 32.98 | 5.74 | 0.97705 |

Table 2 presents the comparative regression performance of the existing models and the proposed HBEL model using the evaluation metrics of MAE, MSE, RMSE, and R²-score. The existing DT-CART achieved an MAE of 10.044, MSE of 138.58, RMSE of 11.77, and an R²-score of 0.035, while RF-CART produced slightly improved results with an MAE of 10.040, MSE of 137.76, RMSE of 11.73, and an R²-score of 0.040. The existing GB-CART demonstrated better regression performance with an MAE of 8.412, MSE of 98.986, RMSE of 9.94, and an R²-score of 0.310, whereas the existing DQN-CART achieved an MAE of 9.9701, MSE of 138.21, RMSE of 11.75, and an R²-score of 0.030. In comparison, the proposed HBEL CART model achieved the best performance with the lowest MAE of 4.854, MSE of 32.98, RMSE of 5.74, and the highest R²-score of 0.97705. These results demonstrate that the proposed HBEL model significantly improves regression

Furthermore, the proposed model achieved a precision of 98.31%, recall of 98.25%, and F1-score of 98.25%, indicating highly balanced and consistent predictive performance. The improvement in all evaluation metrics demonstrates the effectiveness of integrating BiLSTM with PTT for enhanced feature representation and decision-making. Overall, the comparative results confirm that the proposed HBEL model provides superior classification performance and greater prediction reliability than the existing methods.

accuracy and prediction reliability compared to the existing approaches.

5. CONCLUSION AND FUTURE SCOPE

The research presents an effective solution for optimizing multi-tier IoT resource allocation by employing the proposed HBEL model for CART operation. The system efficiently analyzes IoT device parameters, including CPU usage, memory usage, latency, jitter, and execution time, to determine optimal task allocation and predict the corresponding resource allocation percentage. The existing DT-CART, RF-CART, and GB-CART models exhibit comparatively lower predictive performance, with DT and RF showing limited capability in learning complex feature relationships, while GB-CART achieves moderate improvement. The existing DQN-CART delivers better classification performance but exhibits relatively higher prediction errors in regression. In contrast, the proposed HBEL-CART model, which integrates BiLSTM with PTT, effectively

captures bidirectional feature dependencies and enhances prediction through probabilistic decision-making, resulting in superior classification and regression performance. The proposed model achieved a classification accuracy of 98.25%, precision of 98.31%, recall of 98.25%, and F1-score of 98.25%, while attaining an MAE of 4.854, MSE of 32.98, RMSE of 5.74, and an R^2 -score of 0.97705 for regression.

REFERENCES

- [1]. Zhang, H.; Fu, Z.; Shu, K. Recognizing Ping-Pong Motions Using Inertial Data Based on Machine Learning Classification Algorithms. *IEEE Access* 2019, 7, 167055–167064
- [2]. Rozo, J.D.; Bielza, C.; Larrañaga, P. Clustering of Data Streams with Dynamic Gaussian Mixture Models: An IoT Application in Industrial Processes. *IEEE Internet Things J. (IoT)* 2018, 5, 3533–3547.
- [3]. Ahamed, A.; Eltawil, A.; Na, Y.; Tawil, S.E. A Machine Learning Approach for Structural Health Monitoring Using Noisy Data Sets. *IEEE Trans. Autom. Sci. Eng.* 2020, 17, 900–908.
- [4]. Alfian, G.; Syafrudin, M.; Farooq, U.; Maarif, M.; Alex, M.; Fitriyani, N.L.; Lee, J.; Rhee, J. Improving efficiency of RFID-based traceability system for perishable food by utilizing IoT sensors and machine learning model. *Food Control* 2020, 110, 107016.
- [5]. Alam, Shahzeb, Muhammad Atif Ramzan, Muhammad Zubair, Ubaid Ullah, Rab Nawaz, and Rahmat Ullah. "AI-driven resource allocation in cloud computing: a systematic review revealing critical sustainability and evaluation gaps." *Computing* 108, no. 5 (2026): 67.
- [6]. Saurabh, Kumar, Satyasundara Mahapatra, and Manish Madhava Tripathi. "An Adaptive Bandwidth Scheduling Algorithm for IoT Devices Using Machine Learning Techniques." *Franklin Open* (2026): 100492.
- [7]. Wang, Zhifei, Kechao Wang, and Jiarun Wang. "Optimal Federated Learning-based Resource Allocation Scheme for NOMA-enabled Industrial Internet of Things." *IEEE Internet of Things Journal* (2026).
- [8]. Akashwar, G., and M. Balaji. "IoT-Based Data Acquisition and ML Forecasting for Intelligent Control of Solar Energy Microgrids." In *2026 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)*, pp. 1-8. IEEE, 2026.
- [9]. Amodu, Oluwatosin Ahmed, Zurina Mohd Hanapi, Raja Azlina Raja Mahmood, Chedia Jarray, Faten A. Saif, Huda Althumali, Umar Ali Bukar, and Mohammed Sani Adam. "A systematic mapping and review on machine learning for non-terrestrial networks assisted Internet of Things: Enabling technologies." *ICT Express* (2026).
- [10]. Solomon, Michael. "Artificial Intelligence for Intelligent Resource Allocation in Cloud and Edge Computing." *Available at SSRN 6121146* (2026).
- [11]. Tang, Yuzang, and Cheng Zha. "Application of internet of things (IoT) and machine learning (ML) in intelligent logistics supply chain." *Discover Internet of Things* (2026).
- [12]. Sudha, S. P., and J. B. Lorent. "A review on machine learning-based precision agriculture techniques for crop farming monitoring with IOT." *Discover Environment* 4, no. 1 (2026): 10.
- [13]. Lin, YueHong. "Research on the Application of the Internet of Things in Optimizing the Allocation of Ideological and Political Education Resources Using Machine Learning." *International Journal of High Speed Electronics and Systems* 35, no. 02 (2026): 2540160.
- [14]. Alam, MD Tanweer, Voruganti Naresh Kumar, D. Keerthy Vasan, Nandikonda Madhavi, and JL Aldo Stalin. "Combined Q-Learning Methods for Machine Learning-Based Dynamic Resource Alignment In IIOT Networks." In *Adaptive Technologies for Sustainable Growth*, pp. 82-88. CRC Press, 2026.
- [15]. Khan, Muhammad Abdullah, Usman Iqbal, Donghyeon Kim, Jijun Hwang, and Haejoon

Jung. "Optimizing techniques to support the development of Green IoT." In *Design and Analysis of Green and Sustainable IoT Technologies for Future Wireless Communications*, pp. 159-195. Academic Press, 2026.