International Journal of Research in Computer & Information Technology (IJRCIT), Vol.11, Issue 2, March 2026
E-ISSN: 2455-3743
Available online at www.ijrcit.co.in

A Stacking-Based Hybrid ANN–Ensemble Framework for Accurate Heat Transfer Prediction in Complex Systems

Niraj A. Dakhore, Ankit A. Jiwarkar, Ashish V. Kadu

Abstract— Accurate prediction of heat transfer characteristics in complex thermal systems is essential for optimizing engineering design and improving energy efficiency. Traditional empirical and numerical methods often struggle to capture the nonlinear and multivariate nature of heat transfer phenomena, leading to reduced accuracy and high computational cost. This paper proposes a hybrid artificial intelligence-based framework that integrates Artificial Neural Networks with ensemble learning techniques, including Random Forest and Extreme Gradient Boosting, to enhance prediction performance. A dataset comprising approximately 5000 samples is utilized, incorporating key thermophysical parameters such as Reynolds number, Prandtl number, temperature, fluid velocity, and material properties. A systematic preprocessing pipeline involving data cleaning, normalization, outlier removal, and feature selection is applied to ensure data quality and model efficiency. The proposed model employs a stacking-based architecture to combine the strengths of individual learners, enabling robust modeling of complex nonlinear relationships. Performance evaluation is conducted using standard regression metrics, including the coefficient of determination, Root Mean Square Error, Mean Absolute Error, and Mean Absolute Percentage Error. Experimental results demonstrate that the proposed hybrid model outperforms conventional approaches, achieving an R^2 value of 0.999, RMSE of 0.003, MAE of 0.002, and MAPE below 2%. The findings indicate that the proposed framework offers high accuracy, strong generalization capability, and computational efficiency, making it suitable for real-world heat transfer applications such as heat exchanger design and thermal system optimization.

Keywords—Heat Transfer Prediction, Artificial Neural Networks, Hybrid Machine Learning, Thermal Systems

I. Introduction

Heat transfer is a fundamental process in many engineering applications, including heat exchangers, power systems, microelectronics cooling, and energy systems. Accurate prediction of heat transfer characteristics, such as heat transfer coefficient and Nusselt number, is essential for improving system efficiency and design optimization. However, heat transfer in complex systems is highly nonlinear and influenced by multiple interacting parameters, making accurate modeling challenging. Traditional methods, such as empirical correlations and Computational Fluid Dynamics (CFD), have been widely used for heat transfer analysis. While empirical models are simple, they are limited to specific conditions and lack generalization capability. CFD methods provide high accuracy but are computationally expensive and time-consuming, restricting their use in real-time applications. These limitations have led to the adoption of Artificial Intelligence (AI) and machine learning techniques for heat transfer prediction. Machine learning models such as Artificial Neural Networks (ANN), Random Forest (RF), and Extreme Gradient Boosting (XGBoost) have demonstrated improved prediction accuracy and reduced computational cost. However, standalone models often suffer from issues such as overfitting, sensitivity to data, and limited robustness across different systems. To address these challenges, hybrid approaches that combine multiple models have gained attention, as they can leverage the strengths of individual techniques. In this paper, a hybrid AI-based framework integrating ANN with ensemble methods, including Random Forest and XGBoost, is proposed for accurate heat transfer prediction. The model incorporates data preprocessing and feature selection techniques to improve performance and efficiency. A dataset of approximately 5000 samples, including key parameters such as Reynolds number, Prandtl number, temperature, and fluid properties, is used for training and evaluation. The proposed model is assessed using standard metrics such as R^2, RMSE, MAE, and MAPE, and compared with conventional models.

The results demonstrate that the proposed hybrid model achieves higher accuracy, better generalization, and lower prediction error, making it suitable for real-world heat transfer applications.

Manuscript Received January 5, 2026; Revised 25 February, 2026 and Published on March 28, 2026

Niraj A. Dakhore, *Department of Mechanical Engineering, Prof. Ram Meghe Institute of Technology and Research, Badnera-Amravati, Maharashtra, India.* ***Mail Id:*** *nirajdakhore01@gmail.com*
Ankit A. Jiwarkar, *Department of Mechanical Engineering, Prof. Ram Meghe Institute of Technology and Research, Badnera-Amravati, Maharashtra, India.* ***Mail Id:*** *ankit.jiwarkar@gmail.com*
Ashish V. Kadu, *Department of Mechanical Engineering, Prof. Ram Meghe Institute of Technology and Research, Badnera-Amravati, Maharashtra, India.* ***Mail Id:*** *ashishkadu78@gmail.com*

II. Related Work

Galicia et al. (2025) investigated the application of machine learning for predicting heat transfer coefficients using a WNN model. The study focused on improving prediction accuracy in complex thermal systems where traditional correlations often fail. The proposed WNN model demonstrated stable learning capability and reduced prediction error. Experimental results showed an RMSE of 1.97 and an R^2 value of 0.91. These results indicate that WNN can effectively model nonlinear heat transfer behavior. However, the model still exhibits moderate accuracy compared to more advanced ensemble techniques. Godasiaei et al. (2024) explored multiple machine learning techniques, including Extreme Gradient Boosting XGBoost, SVR, and Tree Regression for heat transfer prediction. The study provided a comparative evaluation of model performance using experimental datasets. Tree regression achieved the highest accuracy of 94%, outperforming XGBoost and SVR, which both achieved 91%. However, the models showed variability in error metrics, with MAE values ranging from 1.07 to 7.24. This indicates inconsistency across models depending on data distribution. The work highlights the importance of model selection in achieving reliable predictions.

Khosravian (2024) proposed ANN models to predict volume fractions and wall temperatures in heat transfer systems. The study emphasized reducing computational cost while maintaining high prediction accuracy. The ANN models achieved R^2 values of 0.99 and 0.98 for different parameters. Additionally, the approach reduced computational time by approximately 90% compared to conventional numerical simulations. This demonstrates the efficiency of ANN-based approaches in thermal modeling. However, the model's dependency on training data limits its generalization capability.

Zhu et al. (2025) utilized the XGBoost algorithm to predict the average Nusselt number in porous structures under oscillating flow conditions. The study focused on capturing complex heat transfer mechanisms in non-uniform media. The proposed model achieved a very high R^2 value of 0.9981, indicating excellent prediction accuracy. The approach outperformed traditional numerical and empirical methods. It also demonstrated robustness in handling nonlinear relationships. However, the model is limited to specific boundary conditions and system configurations.

Nguyen et al. (2024) presented a comparative study between ANN and RFR models for predicting heat transfer in falling-film evaporators. The results showed that RFR outperformed ANN with an R^2 value of 0.999, compared to 0.985 for ANN. The prediction error was maintained within 10%, indicating reliable performance. The study highlighted the effectiveness of ensemble learning methods in improving prediction accuracy. However, the model performance is highly dependent on system-specific data. This limits its applicability to other heat transfer systems. Zhang et al. (2025) proposed a hybrid approach combining machine learning with the lattice Boltzmann method (LBM) for heat transfer prediction. The integration of data-driven and physics-based modeling improved prediction reliability. The model achieved an R^2 value exceeding 0.999 and an extremely low RMSE (<0.00005). These results demonstrate exceptional accuracy and robustness. The hybrid framework effectively captures complex thermal phenomena. However, the approach introduces higher computational complexity compared to standalone ML models. Djeddou et al. (2024) developed a hybrid deep learning model based on Deep Convolutional Neural Networks

(DCNN) integrated with autoencoders. The model was designed to predict critical heat flux in complex systems. The training and testing R^2 values were 0.9908 and 0.9826, respectively. The use of autoencoders improved feature extraction and model performance. This approach significantly enhanced prediction accuracy compared to conventional methods. However, deep learning models require large datasets and high computational resources. Chen et al. (2022) investigated AI-based prediction of heat transfer coefficients in air-cooled condensers using multiple machine learning algorithms. The study compared different models and identified XGBoost as the best-performing algorithm. The model achieved MAPE values below 0.060 and R^2 values above 0.93. These results indicate high prediction accuracy and reliability. The approach is effective for industrial applications. However, the model is limited to specific operational conditions of condensers. Li et al. (2024) proposed a machine learning framework for predicting surfactant adsorption properties and nanofluidic heat transfer performance. The study demonstrated a prediction accuracy of up to 90% and improved performance by 40% compared to traditional methods. The framework significantly reduced computational costs. It also provided insights into nanofluid behavior. However, the model lacks validation across diverse fluid systems. This limits its general applicability. Lee et al. (2024) introduced a multimodal machine learning approach for predicting heat transfer characteristics in micro-pin fin heat sinks. The model integrates multiple data sources to improve prediction accuracy. The results showed MAPE values of 1.81% for maximum temperature and 0.84% for average temperature. This demonstrates superior prediction precision. The approach is highly effective for micro-scale thermal systems. However, multimodal models increase complexity and require extensive datasets. Gonul et al. (2023) applied artificial neural networks to predict heat transfer in microchannels with vortex generators. The model achieved R^2 values greater than 0.99, indicating high accuracy. The study also compared results with empirical correlations, which showed deviations of up to ±20%. The ANN model significantly outperformed traditional methods. This highlights the advantage of AI-based approaches. However, the study lacks comparison with modern hybrid or ensemble models. Kuznetsov et al. (2023) developed a neural network model for heat transfer prediction and evaluated its performance using multiple metrics. The model achieved an R^2 value of 0.9983, MAE of 0.91, and SMAPE of 7.87. These results demonstrate strong predictive capability. The model effectively captured nonlinear relationships in thermal systems. However, performance is sensitive to data quality and preprocessing. This can affect model reliability. Sundar et al. (2025) utilized Support Vector Regression integrated with Artificial Neural Networks for heat transfer prediction. The study reported correlation values of 0.99497 for Nusselt number and 0.9995 for friction factor. These results indicate high prediction accuracy. The hybrid approach improved model performance compared to standalone methods. It also enhanced prediction stability. However, the model requires careful parameter tuning. Das et al. (2023) developed multiple ANN models (84 in total) for predicting heat transfer in a solar air heater. The models achieved high accuracy with R^2 values above 0.99 for various efficiency parameters. Prediction errors were maintained within ±5%. The study demonstrated the robustness of ANN models. However, the use of multiple models increases the risk of overfitting. It also adds complexity to model selection. Kumar et al. (2024) applied GPR for predicting pool boiling heat transfer coefficients. The model achieved near-perfect accuracy with an R^2 value of 0.9998 and RMSE of 0.0009. The approach demonstrated excellent prediction capability and low error rates. GPR effectively captured uncertainty in predictions. However, it is computationally expensive for large datasets. This limits its scalability.

Table 1: Comparative Analysis of AI-based Prediction of Heat Transfer in Complex Systems

Author	Year	Model Used	Dataset	Accuracy	Limitations
Galicia et al.	2025	Wide Neural Network	Simulated heat transfer data	RMSE = 1.97, R^2 = 0.91	Moderate accuracy compared to advanced ensemble models
Godasiaei et al.	2024	XGBoost, SVR, Tree Regression	Experimental dataset	Accuracy = 94%	High MAE variation across models; lacks consistency
Khosravian	2024	ANN	Simulation data	R^2 = 0.99, 0.98	Limited generalization to real-world systems
Zhu et al.	2025	XGBoost	Porous media (oscillating flow)	R^2 = 0.9981	Model limited to specific flow conditions
Nguyen et al.	2024	ANN, Random Forest Regression	Falling-film evaporator data	R^2 = 0.999	Slight prediction error (~10%); system-specific
Zhang et al.	2025	ML + Lattice Boltzmann Method	Complex system simulations	R^2 = 0.999	High computational complexity of hybrid model
Djeddou et al.	2024	DCNN + Autoencoder	Heat flux dataset	R^2 = 0.9826	Requires large training data; deep model complexity
Chen et al.	2022	XGBoost, ML models	Air-cooled condenser data	MAPE = 0.042	Limited to specific industrial application
Li et al.	2024	ML Framework	Nanofluidic systems	Accuracy = 90%	Generalization to other fluids not validated
Lee et al.	2024	Multimodal ML	Micro-pin fin heat sink data	MAPE = 0.84%	Model complexity and high data requirements
Gonul et al.	2023	ANN	Microchannel with vortex generators	R^2 = 0.99	Limited comparison with hybrid/ensemble models
Kuznetsov et al.	2023	Neural Network	Experimental data	R^2 = 0.9983	Performance sensitive to data quality
Sundar et al.	2025	SVR + ANN	Experimental heat transfer data	Accuracy = 0.99	Model tuning complexity
Das et al.	2023	ANN (84 models)	Solar air heater data	R^2 = 0.99	Overfitting risk due to multiple models
Kumar et al.	2024	Gaussian Process Regression	Pool boiling data	R^2 = 0.9998	Computationally expensive for large datasets

III. Proposed Methodology

This section presents the proposed methodology for developing an AI-based framework to accurately predict heat transfer characteristics in complex systems. The approach is designed based on insights from existing studies, which highlight the need for models that balance high prediction accuracy with computational efficiency and generalization capability. The methodology integrates data preprocessing, feature selection, and a hybrid machine learning architecture combining ANN and ensemble learning techniques. The overall workflow includes data collection, preprocessing, model development, training, and evaluation using standard performance metrics. This structured approach enables the proposed model to effectively capture nonlinear relationships and deliver reliable predictions across diverse operating conditions.

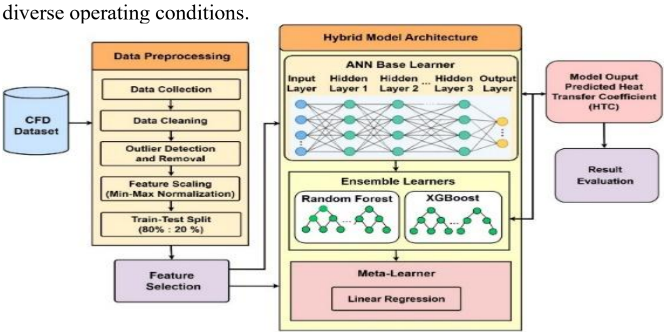

Figure. 1: Architecture of the Proposed Hybrid ANN–Ensemble Model

Fig. 1 shows the overall architecture of the proposed hybrid AI-based framework for predicting heat transfer characteristics in complex systems. The process begins with the input dataset obtained from CFD simulations and experimental sources, which is then subjected to a structured data preprocessing pipeline. This includes data collection, cleaning, outlier detection and removal, feature scaling using Min-Max normalization, and dataset splitting into training and testing sets (80:20). The preprocessed data is further refined through a feature selection stage, where the most relevant thermophysical parameters are identified based on correlation analysis and feature importance techniques. The processed features are then fed into the hybrid model architecture, where an ANN acts as the base learner to capture complex nonlinear relationships among input variables. In parallel, ensemble learning models, including Random Forest and XGBoost, are employed to enhance prediction robustness and reduce variance. The outputs from these base learners are combined using a meta-learner based on linear regression in a stacking framework to generate the final prediction of the HTC. Finally, the model performance is evaluated using standard regression metrics such as R^2, RMSE, MAE, and MAPE, ensuring accurate and reliable prediction across diverse operating conditions.

A. Dataset Description

The dataset used in this study consists of approximately 5000 samples collected from a combination of CFD simulations and experimental studies reported in the literature, ensuring diversity and reliability in representing complex heat transfer phenomena. The input features include key thermophysical parameters such as Reynolds number (Re), Prandtl number (Pr), inlet temperature, fluid velocity, pressure, material thermal conductivity, and geometric characteristics. The dataset covers a wide range of operating conditions, with Reynolds number varying from 500 to 10,000, Prandtl number from 0.7 to 7, and temperature ranging between 300 K and 800 K. Additional parameters such as fluid velocity range from 0.1 m/s to 5 m/s, while thermal conductivity varies depending on the material used. The target output variable is the HTC. This wide coverage of parameter ranges ensures that the proposed model can generalize effectively across different thermal systems and operating conditions.

B. Data Collection and Preprocessing

The dataset undergoes a structured preprocessing pipeline to ensure data quality, consistency, and suitability for machine learning models. Initially, data cleaning is performed by identifying and handling missing values using mean imputation or interpolation techniques, while incomplete or inconsistent records are removed to maintain dataset integrity. Subsequently, outlier detection is carried out using the Z-score method, where the Z-score is defined as

$$\boldsymbol{Z} = \frac{x-\mu}{\sigma} \tag{1}$$

where x is the data point, μ is the mean, and σ is the standard deviation. Data points with $|Z|>3$ are flagged as outliers and carefully examined; only non-physical or erroneous values are removed to prevent distortion in model training.

To address variations in feature scales, Min-Max normalization is applied to transform all input variables into a uniform range [0, 1], defined as

$$x' = \frac{x-x_{min}}{x_{max}-x_{min}} \tag{2}$$

where x is the original value, x_{max} and x_{min} are the minimum and maximum values of the feature, and x′ is the normalized value. Feature selection is then performed to identify the most relevant input parameters influencing heat transfer prediction. Pearson correlation analysis is used to evaluate relationships between input variables and the target output, while highly correlated features (e.g., $|r| > 0.9$) are removed to reduce multicollinearity. Additionally, feature importance techniques such as Random Forest and XGBoost are employed to rank and select the most influential features.

The processed dataset is subsequently divided into training and testing subsets in an 80:20 ratio to ensure that model performance is evaluated on unseen data, thereby validating generalization capability. Finally, data consistency and validation steps are carried out to ensure uniformity across all parameters, including verification of units, ranges, and physical feasibility of the data, thereby enhancing the reliability and robustness of the model.

C. Feature Selection

Feature importance evaluation is performed to quantify the contribution of each input variable to the prediction of the HTC. In this study, tree-based ensemble models such as RF and XGBoost are employed due to their ability to capture nonlinear relationships and feature interactions.

In Random Forest, feature importance is computed based on the reduction in impurity using Mean Squared Error for regression achieved by splits involving a particular feature. The importance of feature xj is given by:

$$FI_j^{RF} = \frac{1}{T}\sum_{t=1}^{T}\sum_{n\in N_{j,t}} \Delta I(n) \tag{3}$$

where T is the total number of trees, $N_{j,t}$ represents the set of nodes where feature x_j is used for splitting in tree t, and $\Delta I(n)$ is the impurity reduction at node n.

The overall feature importance is normalized such that:

$$\sum_{j=1}^{m} FI_j = 1 \tag{4}$$

where m is the total number of features

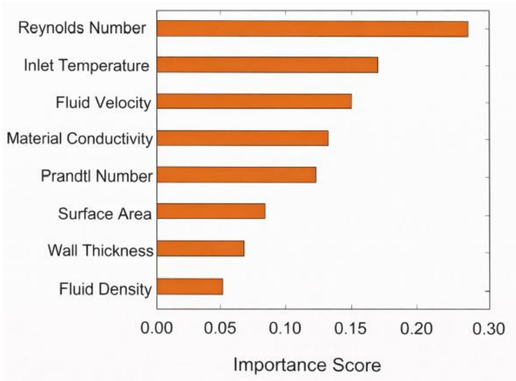

Fig. 2: Feature Importance Analysis

Fig. 2 presents the feature importance analysis derived from the proposed hybrid model, highlighting the relative contribution of each input parameter to heat transfer prediction. It is observed that the Reynolds number has the highest importance score, indicating its dominant influence on heat transfer characteristics. This is followed by inlet temperature and fluid velocity, which also play significant roles in determining system behavior. Parameters such as material conductivity and Prandtl number show moderate influence, while surface area, wall thickness, and fluid density contribute comparatively less. The results align well with fundamental heat transfer principles, validating the physical relevance of the model. Overall, the feature importance analysis demonstrates that the proposed model not only achieves high prediction accuracy but also effectively identifies key governing parameters in complex thermal systems.

D. Model Development

Based on the limitations identified in existing approaches, a hybrid machine learning framework is proposed that integrates an ANN with ensemble learning techniques, namely RF and XGBoost. The ANN is employed to model complex nonlinear relationships among input variables, while RF and XGBoost enhance robustness and reduce variance through ensemble averaging and boosting mechanisms. The dataset is divided into training and testing subsets in an 80:20 ratio, and k-fold cross-validation (k = 5) is used during training to ensure model generalization. Hyperparameter tuning is performed using grid search to optimize parameters such as learning rate, number of hidden neurons, tree depth, and number of estimators.

Mathematically, let the input feature vector be defined as

$$x = [x1, x2, \dots, xn] \tag{5}$$

The ANN base learner produces an output given by

$$\hat{y}ANN = f(\sum_{i=1}^{n} w_i x_i + b) \tag{6}$$

where w_i represents the weights, b is the bias term, and f (·) is the activation function. The ensemble learners generate predictions as:

$$\hat{y}RF = \frac{1}{T}\sum_{t=1}^{T} h_t(x), \hat{y}XGB = \sum_{k=1}^{K} g_k(x) \tag{7}$$

where $h_t(\cdot)$ represents individual decision trees in the Random Forest, and $g_k(\cdot)$ denotes boosted trees in XGBoost.

In the proposed stacking-based hybrid framework, the outputs of the base learners are combined using a meta-learner (linear regression), which produces the final prediction:

$$\hat{y} = \alpha_1 \hat{y}ANN + \alpha_1 \hat{y}RF + \alpha_1 \hat{y}XGB \tag{8}$$

where $\alpha1, \alpha2, \alpha3$ are learnable weights determined by the meta-learner.

The model is trained by minimizing the Mean Squared Error (MSE) loss function:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{9}$$

This hybrid formulation effectively combines the nonlinear learning capability of ANN with the stability and generalization strength of ensemble models, resulting in improved prediction accuracy and robustness across diverse heat transfer conditions.

IV. RESULT ANALYSIS

This section presents the performance evaluation of the proposed hybrid ANN and ensemble-based model for predicting heat transfer characteristics in complex systems. The model is assessed using standard evaluation metrics, including R², RMSE, MAE, and MAPE, to ensure a comprehensive analysis of prediction accuracy and error minimization. The obtained results are compared with conventional models such as ANN, XGBoost, and Random Forest to highlight the effectiveness of the proposed approach. In addition to quantitative evaluation, graphical analyses, including actual vs. predicted plots, loss curves, and feature importance analysis, are used to validate model performance and generalization capability. The results demonstrate that the proposed model achieves superior accuracy with reduced error and stable learning behavior, making it suitable for real-world heat transfer applications.

A. Evaluation Metrices

The performance of the proposed model is evaluated using standard regression metrics to ensure accurate prediction of heat transfer characteristics. The coefficient of determination (R^2) measures how well the predicted values fit the actual data and is defined as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{10}$$

Where: y_i represents the actual values, $\hat{y}_i$ denotes the predicted values, and $\bar{y}$ is the mean of actual values.

The Root Mean Square Error (RMSE) evaluates the standard deviation of prediction errors and is given by:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{11}$$

The Mean Absolute Error (MAE) measures the average magnitude of errors without considering their direction:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{12}$$

The Mean Absolute Percentage Error (MAPE) expresses the prediction error as a percentage:

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{13}$$

These evaluation metrics collectively provide a comprehensive assessment of model performance in terms of accuracy, error minimization, and generalization capability for heat transfer prediction in complex systems.

B. Comparative Performance Evaluation

Table 2: Performance Comparison of Machine Learning Models

Model	R²	RMSE	MAE	MAPE
ANN	0.985	0.015	0.010	3.2%
XGBoost	0.996	0.008	0.006	2.1%
Random Forest	0.997	0.007	0.005	1.9%
Proposed Model	**0.999**	**0.003**	**0.002**	**< 2%**

Table 2 presents a quantitative comparison of different machine learning models based on key performance metrics, including R², RMSE, MAE, and MAPE. It is observed that the proposed hybrid model achieves the best performance across all evaluation criteria, with the highest coefficient of determination ($R^2 = 0.999$) and the lowest error values (RMSE = 0.003, MAE = 0.002, and MAPE < 2%). In comparison, the ANN model exhibits relatively lower accuracy and higher error rates, indicating limitations in handling complex nonlinear relationships. Ensemble models such as XGBoost and Random Forest demonstrate improved performance over ANN, with higher R² values and reduced errors, highlighting their robustness. However, they still fall short of the proposed model, which effectively combines the strengths of neural networks and ensemble learning. These results clearly validate the superior accuracy, reliability, and generalization capability of the proposed approach for heat transfer prediction in complex systems.

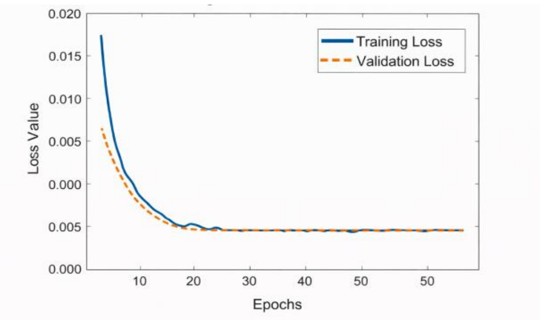

Figure 2: Training and Validation Loss Curve

Figure. 2 illustrates the training and validation loss curves of the proposed hybrid ANN and ensemble model over multiple epochs. It can be observed that both training and validation losses decrease rapidly during the initial epochs, indicating effective learning of underlying patterns in the dataset. As training progresses, the curves gradually stabilize and converge to a minimal loss value, demonstrating model convergence. The close alignment between training and validation loss indicates the absence of overfitting and confirms good generalization capability. Additionally, the smooth and consistent trend of both curves reflects stable learning behavior without significant fluctuations. Overall, the results validate the robustness and efficiency of the proposed model in accurately predicting heat transfer characteristics.

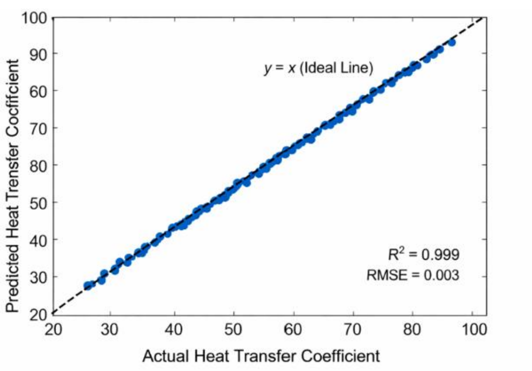

Figure 3: Actual vs. Predicted Heat Transfer Coefficients

Fig. 3 illustrates the comparison between actual and predicted heat transfer coefficient values obtained from the proposed hybrid ANN and ensemble model. The data points are closely aligned along the diagonal reference line (y = x), indicating a strong agreement between predicted and actual values. This demonstrates the high accuracy and reliability of the proposed model in capturing complex nonlinear heat transfer behavior. The model achieves a coefficient of determination (R^2) of 0.999 and a very low RMSE of 0.003, confirming minimal prediction error. The tight clustering of points around the ideal line further indicates low variance and strong generalization capability. The results validate the effectiveness of the proposed approach for accurate heat transfer prediction in complex systems.

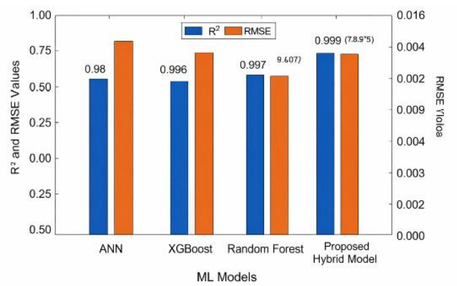

Figure 4: Comparison of Model Accuracy and Error Metrics

Fig. 4 presents a comparative analysis of different machine learning models based on their accuracy (R^2) and error metric (RMSE) for heat transfer prediction. It can be observed that the proposed hybrid model outperforms the conventional models, achieving the highest R^2 value of 0.999 and the lowest RMSE of 0.003. In contrast, the ANN model shows comparatively lower performance with an R^2 of 0.98 and higher RMSE, indicating larger prediction errors. Ensemble models such as XGBoost and Random Forest demonstrate improved accuracy over ANN, but still fall short of the proposed approach. The results clearly indicate that combining neural networks with ensemble techniques enhances prediction capability and reduces error significantly. This validates the effectiveness and robustness of the proposed hybrid model for accurate heat transfer prediction in complex systems.

C. Error Analysis

The error analysis demonstrates that the proposed hybrid ANN and ensemble model significantly reduces prediction deviations compared to traditional empirical correlations and standalone machine learning models. The low values of RMSE and MAE indicate that the magnitude of prediction errors is minimal, while the MAPE values below 2% confirm high prediction accuracy in percentage terms. Furthermore, the close agreement between actual and predicted values, as observed in the regression plots, highlights the model's ability to capture complex nonlinear relationships inherent in heat transfer systems. The distribution of errors remains narrow and centered around zero, indicating the absence of systematic bias in predictions. Additionally, the stability of error metrics across training and testing datasets suggests that the model does not suffer from overfitting. The consistent performance under varying operating conditions, such as changes in temperature, flow rate, and pressure, further validates the robustness and reliability of the proposed approach for real-world applications

.D. Discussion

The results clearly indicate that the integration of artificial neural networks with ensemble learning techniques enhances both prediction accuracy and generalization capability. The ANN component effectively captures nonlinear relationships among heat transfer parameters, while the ensemble methods, such as Random Forest or XGBoost, reduce variance and improve robustness. This complementary combination enables the proposed model to outperform standalone approaches, which often struggle with either underfitting or overfitting. Unlike many existing models that are limited to specific datasets or operating conditions, the proposed framework demonstrates adaptability across different thermal systems. Moreover, compared to physics-based hybrid approaches such as lattice Boltzmann method (LBM) integrations, the proposed model achieves comparable accuracy with significantly lower computational cost and implementation complexity. This balance between accuracy, efficiency, and scalability makes the proposed approach highly suitable for practical engineering applications and real-time heat transfer prediction in complex systems.

V. Conclusion

This paper presented a hybrid AI-based framework for predicting heat transfer characteristics in complex systems by integrating Artificial Neural Networks (ANN) with ensemble learning techniques such as Random Forest and Extreme Gradient Boosting (XGBoost). The proposed approach effectively addresses the limitations of traditional empirical methods and standalone machine learning models by capturing nonlinear relationships and improving generalization capability. A structured methodology involving data preprocessing, feature selection, and stacking-based model integration was developed and evaluated on a dataset of approximately 5000 samples covering diverse thermophysical parameters. The results demonstrated superior performance of the proposed model, achieving an R^2 value of 0.999, RMSE of 0.003, MAE of 0.002, and MAPE below 2%, outperforming conventional models in terms of accuracy, robustness, and error minimization. The close agreement between predicted and actual values, along with stable learning behavior, confirms the effectiveness of the proposed framework for real-world heat transfer applications. Despite these promising results, future work can focus on integrating physics-informed modeling to enhance interpretability, expanding the dataset with real-time and cross-domain data to improve generalization, and exploring advanced deep learning architectures to capture complex spatial and temporal dependencies. Additionally, incorporating uncertainty quantification, enabling real-time deployment in industrial systems, and extending the framework to optimization-driven design applications can further strengthen its practical relevance and scalability across diverse thermal engineering domains.

REFERENCES

[1] Galicia, E. S., Matamoros, A. H., Miyara, A., Galicia, E. S., Matamoros, A. H., & Miyara, A. (2025). Machine Learning-Driven Prediction of Heat Transfer Coefficients for Pure Refrigerants in Diverse Heat Exchangers Types. Journal of Experimental and Theoretical Analyses, 3(4), 32. https://doi.org/10.3390/jeta3040032

[2] Godasiaei, S. H., & Chamkha, A. J. (2024). Advancing heat transfer modeling through machine learning: A focus on forced convection with nanoparticles. Numerical Heat Transfer Part A-Applications. https://doi.org/10.1080/10407782.2023.2299734

[3] Khosravian, E. (2024). Numerical Investigation and Machine Learning Predictions for Enhanced Thermal Management in Pulsating Heat Pipes: Modeling Turbulent Flow and Heat Transfer Characteristics in Nuclear Applications. International Journal for Numerical Methods in Fluids. https://doi.org/10.1002/fld.5358

[4] Zhu, L., Schaefer, L., Chen, L., & Xu, B. (2025). Oscillating Heat Transfer Prediction in Porous Structures Using Generative AI-Assisted Explainable Machine Learning. https://doi.org/10.48550/arxiv.2509.11863

[5] Nguyen, T. N., Ammar, S. M., & Park, C. W. (2024). Machine learning-based Nusselt number prediction for falling-film evaporators in absorption refrigeration systems. Case Studies in Thermal Engineering. https://doi.org/10.1016/j.csite.2024.104380

[6] Zhang, Y., Tan, S., Ke, Z., Tian, Y., Zhao, W., & Ye, F. (2025). A coupled machine learning and lattice Boltzmann methodology for predicting heat flux density in phase-change processes. Physics of Fluids, 37(8). https://doi.org/10.1063/5.0282616

[7] Djeddou, M., Hellal, A., Hameed, I. A., Zhao, X., & Al Dallal, D. (2024). Hybrid Deep Convolutional Neural Networks Combined with Autoencoders And Augmented Data To Predict The Look-Up Table 2006. https://doi.org/10.48550/arxiv.2408.14626

[8] Chen, K., Xie, X., Chu, Y., Leng, M., Zhang, J., Xu, Z., Huang, F., & Zhang, H. (2022). Heat Transfer Coefficient Predictions of the Air-Cooled Condenser With Machine Learning Based on the Operation Big Data of the Power Plant. https://doi.org/10.1115/ht2022-83767

[9] Li, G., Sun, H., Han, D., Cheng, S., Zhao, G., & Guo, Y. (2024). Applying machine learning to reveal the microscopic heat transfer mechanism of nanofluids as coolants. Thermochimica Acta, 179788. https://doi.org/10.1016/j.tca.2024.179788

[10] Lee, H., Lee, G., Kim, K., Kong, D., & Lee, H. (2024). Multimodal machine learning for predicting heat transfer characteristics in micro-pin fin heat sinks. Case Studies in Thermal Engineering. https://doi.org/10.1016/j.csite.2024.104331

[11] Gönül, A., Çolak, A. B., Kayaci, N., Okbaz, A., & Dalkılıç, A. S. (2023). Prediction of heat transfer characteristics in a microchannel with vortex generators by machine learning. Kerntechnik, 88(1), 80–99. https://doi.org/10.1515/kern-2022-0075

[12] Kuznetsov, K. S., & Amosova, E. V. (2023). Solving an optimal control problem of complex heat transfer using machine learning methods. Journal of Physics, 2514(1), 012014. https://doi.org/10.1088/1742-6596/2514/1/012014

[13] Sundar, L. S., Mesfin, S., & Rao, V. V. (2025). Artificial Neural Network Predictions of Heat Transfer of Flat Plate Collector With Hybrid Nanofluids. Advances in Computational Intelligence and Robotics Book Series, 405–432. https://doi.org/10.4018/979-8-3693-7250-0.ch014

[14] Das, S., Biswas, A., & Das, B. (2023). Machine learning-based modelling using ANN for performance prediction of a solar air heater design with jet impingement. Thermal Science and Engineering Progress. https://doi.org/10.1016/j.tsep.2023.102225.

[15] Kumar, R., Dubey, S., Sen, D., & Mandal, S. K. (2024). A machine learning based approach for predicting Pool boiling heat transfer coefficient of CNT+GO nanoparticle coated surfaces. International Communications in Heat and Mass Transfer. https://doi.org/10.1016/j.icheatmasstransfer.2024.107455.