



ABSTRACTS: VOLUME 8, SPECIAL ISSUE {8th Undergraduate Conference}

PREDICTION OF NANOFLUID SPECIFIC HEAT CAPACITY USING SUPERVISED REGRESSION MODELS

Yomna Abu Amra¹, Ashraf Y. Maghari¹

Supervisor: Dr. Ashraf Y. Maghari¹

¹ Department of Information Technology Systems, Faculty of Information Technology, Islamic University of Gaza, Palestine

Abstract

Background: Accurate prediction of the specific heat capacity (SHC) of nanofluids is essential for improving the design and performance of thermal management, heat transfer, and energy systems. Traditional experimental methods for measuring SHC are often costly, time-consuming, and prone to uncertainties caused by factors such as particle agglomeration, sedimentation, and temperature sensitivity. Therefore, reliable and efficient predictive approaches are needed. Recent advances in machine learning have shown strong potential in modeling complex thermophysical properties of nanofluids and providing fast, cost-effective alternatives to experimental measurements.

Objectives: This study aimed to develop and evaluate supervised machine learning regression models for predicting the specific heat capacity (SHC) of nanofluids. The models were designed to use key thermophysical input parameters, including nanoparticle type, base fluid, nanoparticle volume fraction, and base fluid temperature, to provide an accurate, reliable, and practical predictive tool for engineering and energy applications.

Methods: A dataset of 517 nanofluid records was collected from the Kaggle platform and preprocessed for analysis. Each record included four input features: nanoparticle type, base fluid, nanoparticle volume fraction, and base fluid temperature, while SHC (J/kg·K) was used as the target variable. Categorical variables were encoded using one-hot encoding, and correlation analysis was conducted to identify weak or redundant features. The dataset was split into 70% training and 30% testing sets. Four supervised regression models were developed and compared: Gradient Boosting, XGBoost, AdaBoost, and Decision Tree Regressor. Model performance was evaluated using Mean Squared Error (MSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2). Additionally, 5-fold cross-validation was applied to assess model robustness and reduce overfitting.

Results: The results showed that the Gradient Boosting model achieved the highest predictive performance among all models. On the testing dataset, it obtained an R^2 score of 0.9960 (99.60%), MSE of 2189.93, and



MAE of 27.53. XGBoost followed with an R^2 of 0.9743 (97.43%), AdaBoost achieved 0.9707 (97.07%), and the Decision Tree Regressor achieved 0.8673 (86.73%). Predicted SHC values in representative test scenarios closely matched the original dataset values, confirming the practical applicability of the models. Furthermore, 5-fold cross-validation demonstrated consistent and reliable performance, particularly for Gradient Boosting and XGBoost.

Conclusions: This study demonstrates that supervised machine learning regression models can effectively predict the specific heat capacity of nanofluids with high accuracy. Among the evaluated models, Gradient Boosting proved to be the most accurate and reliable, making it a strong alternative to traditional experimental methods. These findings highlight the potential of machine learning to support faster, more economical, and more efficient design and optimization of advanced thermal and energy systems. Future work may incorporate explainable artificial intelligence techniques, such as SHAP or LIME, to enhance model interpretability and better understand feature importance.

Keywords: Nanofluids; Specific heat capacity; Regression; Gradient boosting; XGBoost.