# Using Artificial Neural Networks with GridSearchCV for Predicting Indoor Temperature in a Smart Home

**Talal Alshammari**
Department of Information and Computer Science, College of Computer Science and Engineering, University of Ha'il, Saudi Arabia
talal.alshammari@uoh.edu.sa

*Received: 4 February 2024 | Revised: 16 February 2024 | Accepted: 19 February 2024*

 *DOI: https://doi.org/10.48084/etasr.7008*

## ABSTRACT

**The acceleration of house technology via the use of mobile phones has made it easier to control houses, where occupants (especially older people) spend most of their time. The climate of Saudi Arabia, especially in the northern area, is too hot during summer and cold during winter. Control of the indoor environment in a smart home is a preferable choice that can reduce power consumption to operate heating, ventilation, and air-conditioning. Machine learning algorithms have been used to predict physical variables of indoor environment, such as temperature and humidity. The model can be trained, learn, and make predictions using historical data. Machine learning techniques can automate temperature monitoring and control. This paper proposes an algorithm that combines Artificial Neural Networks (ANNs) and GridSearchCV to predict physical variables in indoor environments in Saudi Arabia. GridSearchCV was utilized to tune the parameters of the machine learning algorithm. The assessment of the proposed algorithm involved its performance comparison to state-of-the-art machine learning algorithms. A real-world dataset was generated to estimate the performance of the considered algorithms. The room data were collected every 5 min for 31 days during July 2022. The dataset contains 6 columns and 8,910 records from 6 sensors (timestamps, light, temperature, humidity, pressure, and altitude). Random Forest (RF), Decision Tree (DT), and ANN methods were compared with the proposed algorithm. The RF had the highest $R^2$ value of 0.84 and the lowest Mean Square Error (MSE) of 0.43. The DT achieved an $R^2$ score of 0.78, while the ANN achieved $R^2$ score of 0.61, MSE of 1.04, and Mean Absolute Error (MAE) of 0.75. The proposed algorithm achieved an $R^2$ of 0.69, MSE of 0.87, and MAE of 0.67.**

***Keywords-artificial neural networks; deep learning; machine learning algorithms; smart home***

## I. INTRODUCTION

Older people are more susceptible to heat-related illnesses, such as heat exhaustion, heat stroke, and heat cramps. Most older adults take medication that make them more sensitive to heat [1]. Clinical reports have noted that people over 60 have a higher potential to experience heat-related illnesses [2]. Moreover, by 2030, more than 60% of the world's population will live in urban environments, and will be exposed to more increased temperatures than rural areas [3]. According to the statistics, the energy consumption of buildings is around 40% of the total global energy consumption, and buildings emit more than 36% of the total global emissions contributing to carbon dioxide emissions [4, 5]. Consequently, optimal energy use can facilitate the reduction of energy consumption and $CO_2$ emissions [5]. In building technology, the essential goals are reducing energy consumption in buildings and developing a quality indoor environment [10, 11]. Physical variable quality in indoor environments is essential, especially for older people and patients. Hospitals, homes, and healthcare facilities require sufficient monitoring and control of the indoor environment [6, 7]. Occupancy prediction in indoor environments is essential to reduce energy consumption and maintain occupant comfort. Moreover, occupant behavior is important in physical variables, such as temperature, lighting, heating, and air-conditioning [8, 9].

Machine learning algorithms are essential techniques for efficaciously dealing with numerous data. Different applications can predict load forecasting, solar energy, wind energy, and other variables [12, 13]. Even though elevated temperatures are potentially harmful, to the best of our knowledge, there is a lack of an accurate temperature prediction model for older people. For temperature prediction in healthcare, machine learning algorithms are crucial in controlling and monitoring temperatures in smart home environment and can improve accuracy and efficiency in temperature prediction. There are many advantages in the use of machine learning algorithms, such as:

- Recognizing the home temperature for older adults: accurately predicting temperature variations in homes can help the timely intervention of healthcare providers.

- Reducing physical visits: Fewer visits from healthcare providers to check the temperature in the home are needed. In smart homes, smart sensors produce massive sensory data that assists stakeholders to make decisions.

Numerical weather prediction is important to provide vital knowledge for inhabitants and healthcare providers [13]. Therefore, machine learning algorithms have been used to rectify weather forecasting shortcomings with imperfect physical parameterizations [13]. However, measuring the indoor environment for smart homes is challenging. Traditional measurement solutions are insufficient because internal measurements vary and require a massive volume of data for long periods [6, 14]. Internet of Things (IoT) applications provide real-time measurements, and each sensor generates numerous data to apply prediction models to reduce energy consumption and improve environmental quality [6, 14]. Machine learning techniques can predict the indoor environment to control and monitor the temperature and humidity variables and the lighting and Heating, Ventilation, and Air-Conditioning (HVAC) systems. In addition, machine learning techniques are applied to recognize the inhabitants' daily living habits in an indoor environment, monitor health and wellness (especially in older people), and give recommendations for healthcare providers to make decisions. To the best of our knowledge, there is a shortage of investigations and assessments of machine learning algorithms in the scope of indoor environment. Although existing studies on indoor temperature prediction have made good progress, there are many challenges and constraints, such as ensuring data and obtaining the highest performance accuracy [15].

A comprehensive review of the use of ANNs for air temperature prediction is presented in [16]. Different types of ANNS are introduced, such as Recurrent Neural Networks (RNNs) and Long-Short Term Memory (LSTM) Networks. ANNs have the ability to learn complex non-linear relationships from the input data set. The conventional prediction models and their limitations when using ANNs for meteorological data are pointed out. Authors in [17] used ANNs to predict the temperatures inside the heated foil tunnel. The air temperature inside the tunnel plays an important role in managing the heating system, improving crop yield, and reducing energy consumption. The efficiency of ANN models is discussed in forecasting temperature changes inside the heated foil tunnel and the optimized conditions in different industries depend on the monitored environment. Root Mean Square Error (RMSE) was used to evaluate the performance of the ANN model. Authors in [18] proposed a probabilistic ANN that was a combination of yield prediction and optimization, using the hybrid Moth Flame Optimization algorithm with a machine learning algorithm for crop recommendations and yield predictions. The utilized dataset collected information on weather, rain and fertilizer in India. They achieved a prediction accuracy of 99.67% and an R2 value of 98.82%. In [19], DT was applied to develop an automatic irrigation system. The authors installed many sensors under the roots to measure the moisture content in the soil. The sensors recorded the temperature and humidity that were done every 10 min to control the automatic irrigation based on the humidity level to save irrigation water on the agricultural field. The result achieved prediction accuracy that approached 97.86%. Authors in [20] used ANN algorithms to predict maize yield. Six datasets were used to validate the proposed algorithm. The datasets comprise of features of climate, soil water balance and agricultural features of maize for two crop years. The ANN models were able to capture the non-linear effects of the datasets. However, the amount of data for the agricultural environment was limited. RMSE, MSE, and $R^2$ were used to calculate performance of ANN models for predicting maize productivity. Authors in [21] used deep learning to assess the energy consumption of buildings. They pointed out that prediction is crucial in estimating the energy consumption of buildings to save energy. A benchmark dataset was used to estimate their models. The dataset was generated from the electricity consumption of a single occupant measured every minute. FCRBM achieved the best prediction compared to the other machine learning algorithms. Their article gives an overview of different machine learning algorithms that are used to predict air temperature. They show that the deep learning approaches have a lower MSE compared to conventional ANNs. Authors in [22] emphasized the importance of feature selection, data pre-processing and algorithm evaluation in temperature prediction. They used ANN algorithms to predict the air temperature in August in the southern part of the Iberian Peninsula. Authors in [23] used the ERA5 Reanalysis dataset, which integrates monitoring data with numerical weather forecasting techniques to capture temporal and spatial patterns of temperature. The results show that the XAI models are able to efficiently predict air temperature one month ahead with reasonable performance and interpretability.

The main contributions of this paper are: a review of the previous studies on indoor temperature prediction and a comprehensive evaluation of the state-of-the-art machine learning algorithms accompanied with a comparison with the proposed algorithm on the prediction of indoor temperature in a smart home domain based on certain evaluation metrics.

## II. METHODOLOGY

The input data are preprocessed (cleaned) and missing data values are filled in. Real datasets must be used to evaluate the performance of the machine learning algorithms. Only a few real datasets for indoor environments are publicly available. Datasets representing indoor environments with occupancy performance metrics must be generated to assess the models. Predicting indoor environments involves feature importance scores and machine learning algorithms.

## III. REAL INDOOR ENVIRONMENTAL DATASET

The dataset includes measurements of an indoor environment from 6 sensors (timestamp, light, temperature, humidity, pressure, and altitude) from July 2022, consisting of 8,910 real-time records [24]. Samples were captured every 5 min. Table I shows a sample of the dataset.

## IV. EXPERIMENT DESIGN

This section proposes an estimation algorithm for the prediction of temperature. The methodology involves two approaches. In the first approach, the models learn from data without tuning and optimizing the parameters, as in the second approach. The dataset was divided into 80% training and 20% testing subsets. All the input features were fed to the machine learning algorithms with the exception of the timestamp column, because some of the models were not able to handle this data type. A diagram of the proposed model is shown in Figure 1. The proposed model was trained and tested using the optimization method. GridSearchCV was used for finding the best hyperparameters. The GridSearchCV and ANN approach was implemented to improve the performance of the model and apply feature selection to the training and testing sets. The proposed model learns from the data using hyperparameters in the training phase and performs predictions using the trained GridSearch with ANNs. To assess the prediction of the performance of the model, GridSearchCV was employed from Sklearn to carry out a grid search to find the best combination of hyperparameters and estimate classifier's performance using cross-validation.

TABLE I. A DATASET SAMPLE

| Time | Light | Temperature | Humidity | Pressure | Attitude |
|---|---|---|---|---|---|
| 12:02:08 | 357 | 37.5 | 11 | 89535 | 1031.2 |
| 12:07:09 | 352 | 37.6 | 11 | 89536 | 1031.1 |
| 12:12:10 | 349 | 37.6 | 11 | 89532 | 1031.47 |
| 12:33:54 | 339 | 37.7 | 11 | 89510 | 1033.5 |
| 12:38:55 | 336 | 37.7 | 11 | 89512 | 1033.31 |
| 12:43:56 | 330 | 37.7 | 11 | 89507 | 1033.77 |
| 12:48:57 | 325 | 37.7 | 11 | 89501 | 1034.33 |
| 12:53:58 | 322 | 37.9 | 11 | 89502 | 1034.23 |
| 12:58:59 | 315 | 37.9 | 11 | 89499 | 1034.51 |
| 13:04:01 | 312 | 37.9 | 11 | 89492 | 1035.15 |
| 13:09:02 | 304 | 38 | 12 | 89495 | 1034.88 |
| 13:14:03 | 300 | 38 | 12 | 89487 | 1035.61 |

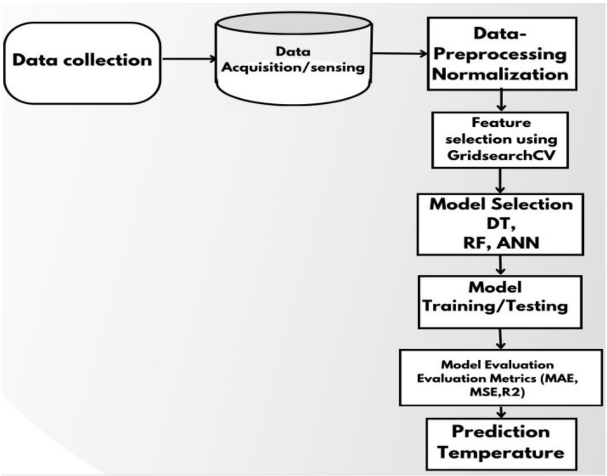


Fig. 1. The proposed optimized model for temperature prediction.

This paper used the Keras library to create a sequential model that illustrates fully-connected layers in the ANN. The model was compiled with the MSE loss function and the Adam optimizer. The model learned from the training data and used GridSearchCV to optimize the parameters. The model was validated with 5-fold cross-validation on the predictions of the testing data. MSE, MAE, and $R^2$ were used as evaluation metrics. Finally, the loss curve of the model during the training phase was obtained and a comparison between the predicted and the actual values was made.

## V. PERFORMANCE METRICS

The MAE between the predictions and the target values is used in regression analysis to measure the model prediction accuracy. The closer the MAE is to zero, the more accurate the predictions and the nearer they are to the target values. MSE is the average of the squared error difference between the predicted and target values. The closer the MSE is to zero, the more accurate the predictions are.

$R^2$ (or the coefficient of determination) measures the estimated value accuracy calculated in a regression model, determining the ratio of variance in the dependent variable that the independent variable can predict. The $R^2$ value is always between 0 and 1. An $R^2$ of 1 indicates that the regression predictions are appropriate for the data, whereas $R^2$ values closer to zero indicate a greater spread [25].

### A. *Decision Tree (DT) Algorithm*

The DT regression method is a supervised learning algorithm for classification and regression problems. It is a nonparametric learning algorithm that predicts continuous values instead of categorical target variables. The tree is structured by recursively dividing each node's data into smaller parts. It aims to reduce the MSE for each child node. The DT model can deal with nonlinear relationships between features and labels and uses categorical and numerical value datasets. However, DTs can be prone to overfitting and cannot handle missing values. Thus, the dataset must be prepared. The tree has nodes linked by edges, and one root node. Each node is connected to a parent node. Figure 2 shows the tree data structure in Python. This method assigns nodes to a root and adds smaller subsets as child nodes [26, 27].

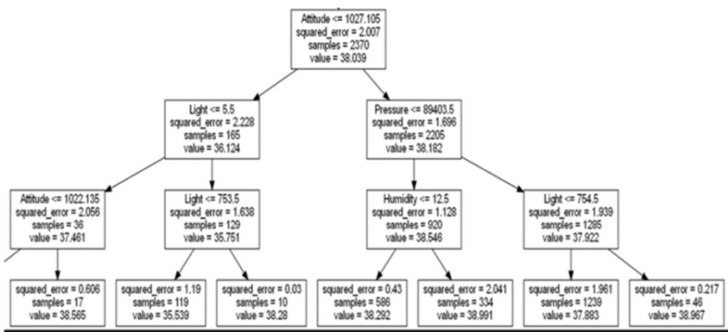


Fig. 2. Export graphviz flow chart.

### B. *Random Forest (RF)*

The RF regression method is a supervised learning algorithm used in regression problems. It is an ensemble of DTs to predict a continuous value. Multiple trees form an RF using a random subset of data and features, and the trees are generated from the samples of features. At each node, a sample of features is selected for splitting, assisting in decreasing overfitting [27, 28].

### C. *Artificial Neutral Network*

An ANN is a feedforward system for classification and regression problems inspired by the human brain. Every node is a neuron. ANNs consist of interconnected neurons arranged in layers, which include several hidden layers between the input and the output. Each layer is completely connected to the following layer by weights, and the output is passed to an activation function. Regression ANNs predict a dependent variable value based on the input characteristics (independent variables) [29].

### *D. GridSearchCV*

GridSearchCV (Grid Search Cross-Validation) is a technique used to tune hyperparameters and model selection [30]. It automates the process to find an optimal set of parameters and avoid overfitting. The proposed model learns from the data in the training phase and the prediction is performed using GridSearchCV with the proposed model. The hyperparameters used in ANNs are batch_size, epochs, hidden_layers, learning_rate and units.

## VI. RESULTS AND DISSCUSION

This study demonstrates the effectiveness of the proposed model in predicting the temperature indoor environment and highlights its potential for optimizing states in diverse indoor environments in a smart home. The combination of ANN and GridSearcCVh in the proposed model showed good error evaluation, which produced the lowest MSE in the considered dataset. The proposed model obtained optimal hyperparameters determined through GridSearchCV and k-fold cross-validation in the dataset. At first, experiments were performed without the use of Grid Search. The accuracy of the machine learning algorithms was evaluated using MAE, MSE, and $R^2$ on the dataset and was determined during the training phase (Table II). Table III shows the results achieved by 5-fold cross-validation. Table IV shows the results revealing the competitive performance of the proposed model and the state-of-the-art machine learning algorithms during the testing phase. The proposed model obtained lower error than the ANN model with normal parameters. The Grid Search was performed for optimizing the proposed algorithm and various parameter settings were used, as shown in Table V. The proposed model's performance had an MSE of roughly 0.87 for testing datasets, MAE of is around 0.67, and $R^2$ of approximately 0.69, demonstrating its ability to predict temperature accurately. Figures 3-6 show the cross-validated predictions during training and the predicted values versus the actual temperature values.

TABLE II. PERFORMANCE RESULTS OF THE EVALUATED MODELS DURING THE TRAINING PHASE

| Model | MSE | MAE | $R^2$ |
|---|---|---|---|
| RF | 0.1538 | 0.2058 | 0.9425 |
| DT | 0.1157 | 0.1206 | 0.9568 |
| ANN | 1.0026 | 0.7425 | 0.6256 |
| Proposed | 0.8937 | 0.6979 | 0.6663 |

TABLE III. PERFORMANCE RESULTS OF THE EVALUATED MODELS BY 5-FOLD CROSS-VALIDATION

| Model | MSE | MAE | $R^2$ |
|---|---|---|---|
| RF | 0.4696 | 0.3865 | 0.8254 |
| DT | 0.6165 | 0.3982 | 0.7709 |
| ANN | 1.0739 | 0.7494 | 0.5885 |
| Proposed | 1.0391 | 0.7451 | 0.6141 |

TABLE IV. RESULTS OF THE PERFORMANCE OF ALL EVALUATED MODELS BY TESTING PHASE.

| Model | MSE | MAE | $R^2$ |
|---|---|---|---|
| RF | 0.4398 | 0.3819 | 0.8404 |
| DT | 0.5970 | 0.3961 | 0.7834 |
| ANN | 1.0496 | 0.7523 | 0.6193 |
| Proposed | 0.87075 | 0.6789 | 0.6913 |

TABLE V. GRIDSEARCHCV SETTINGS USED

| Parameter | Value |
|---|---|
| batch_size | 16 |
| epochs | 100 |
| hidden_layers | 3 |
| learning_rate | 0.01 |
| units | 3 |

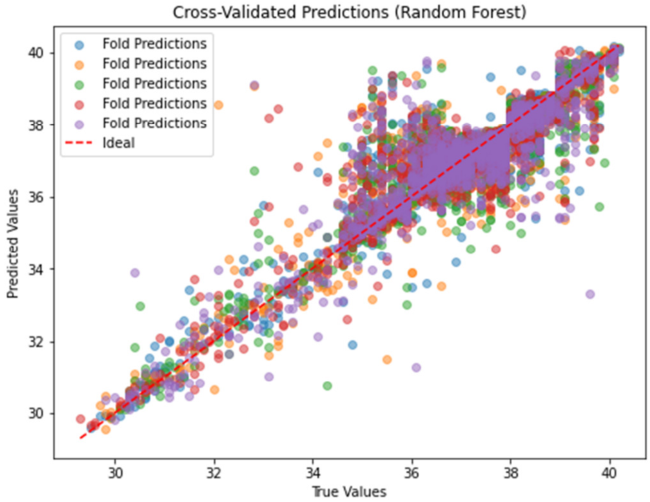


Fig. 3. Cross-validated prediction (RF).

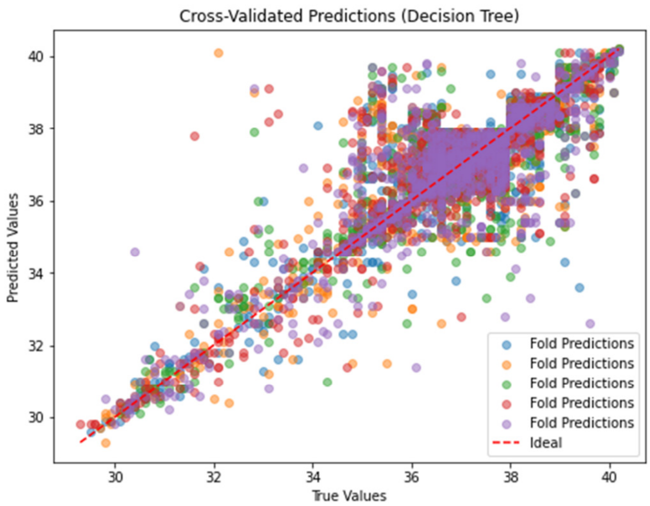


Fig. 4. Cross-validated rediction (DT).

Figures 7-10 show the actual against the predicted temperature during the testing phase. The plot of the training and validation loss obtained during training is shown in Figure 11.

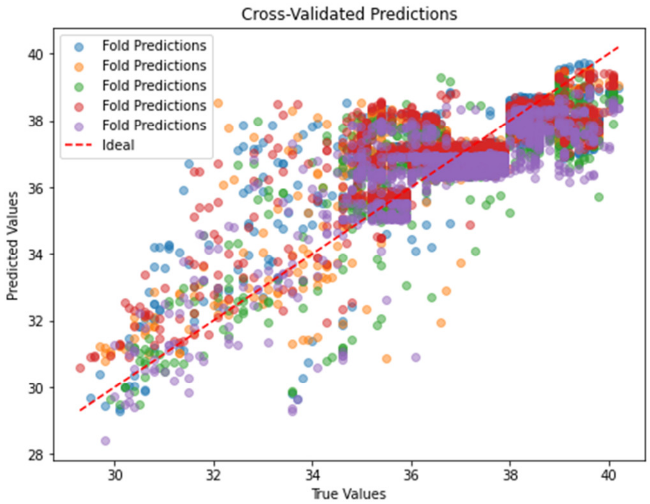


Fig. 5. Cross-validated prediction (ANN).

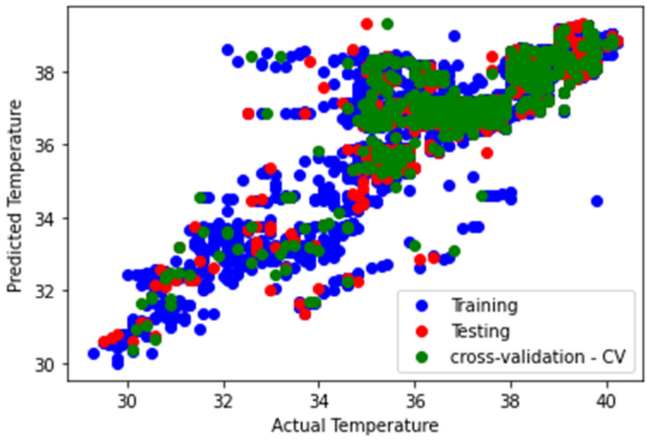


Fig. 6. Cross-validated prediction (proposed model).

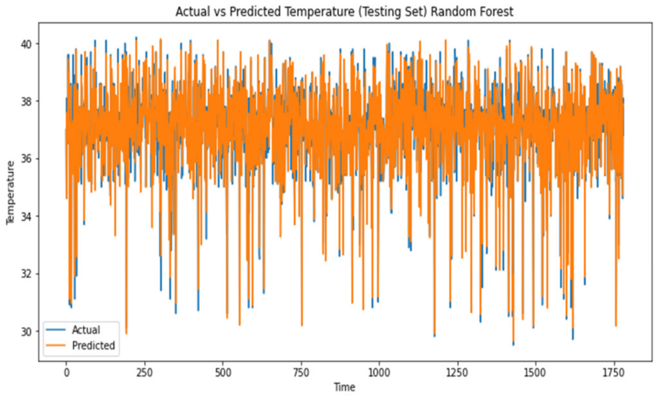


Fig. 7. Actual vs predicted temperature (RF).

## VII. CONCLUSION

This paper presents a useful application for predicting the indoor temperature in a smart home. This study proposes an optimized machine learning algorithm to predict the temperature in the home environment. Various machine learning algorithms were evaluated on a real-world dataset generated from a smart room. Learning from historical indoor physical measurements and predicting future temperature is crucial for smart homes.

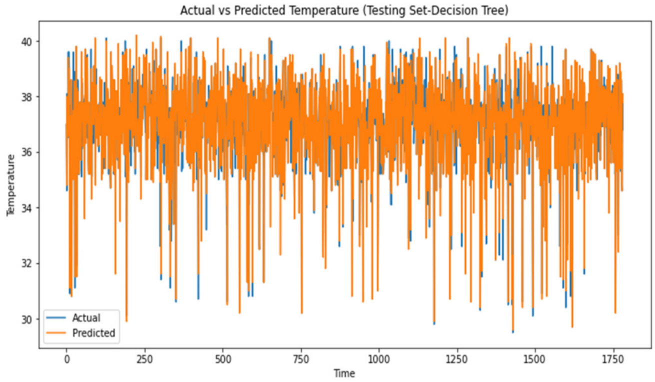


Fig. 8. Actual vs predicted temperature (DT).

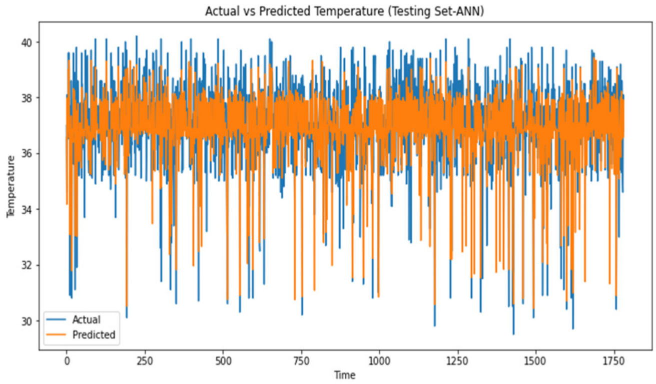


Fig. 9. Actual vs predicted temperature (ANN).

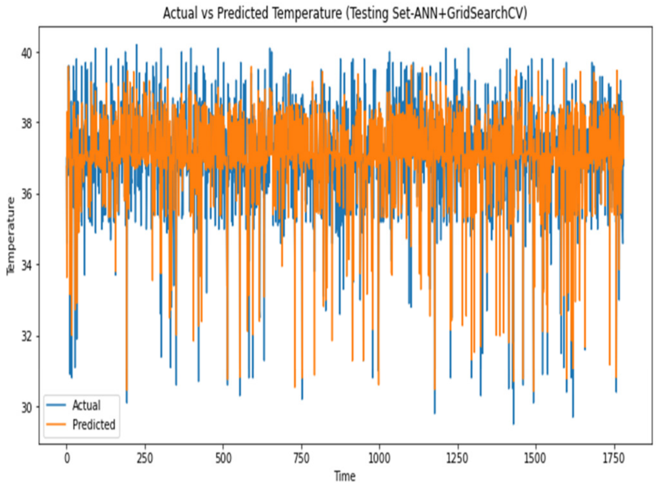


Fig. 10. Actual vs predicted temperature (proposed model).

The considered algorithms were evaluated with $R^2$, MSE, and MAE metrics, including the proposed algorithm against RF, DT, and ANN. RF achieved the highest $R^2$ score of 0.84

whereas the proposed algorithm obtained an $R^2$ of 0.69, MSE of 0.87, and MAE of 0.67. The ANN algorithm achieved an $R^2$ score of 0.61, MSE of 1.04, and MAE of 0.75. However, the results showed that the proposed algorithm obtained better result than the unoptimized ANN. Future work will focus on using reinforcement learning for the prediction of the indoor temperature.

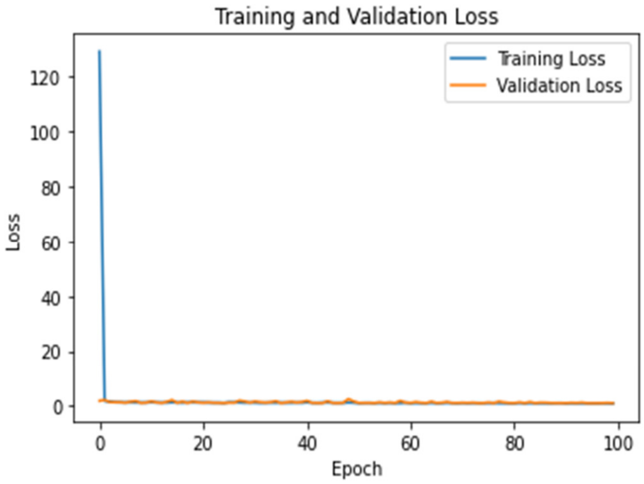


Fig. 11. Training and validation loss (proposed model).

## REFERENCES


[1] R. K. Nath and H. Thapliyal, "Wearable Health Monitoring System for Older Adults in a Smart Home Environment," in *2021 IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*, Tampa, FL, USA, Jul. 2021, pp. 390–395, https://doi.org/10.1109/ISVLSI51109.2021.00077.

[2] B. N. Balmain, S. Sabapathy, M. Louis, and N. R. Morris, "Aging and Thermoregulatory Control: The Clinical Implications of Exercising under Heat Stress in Older Individuals," *BioMed Research International*, vol. 2018, 2018, Art. no. 8306154, https://doi.org/10.1155/2018/8306154.

[3] S. Vulova, F. Meier, D. Fenner, H. Nouri, and B. Kleinschmit, "Summer Nights in Berlin, Germany: Modeling Air Temperature Spatially With Remote Sensing, Crowdsourced Weather Data, and Machine Learning," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 5074–5087, 2020, https://doi.org/10.1109/JSTARS.2020.3019696.

[4] Z. Li, P. Wang, J. Zhang, and S. Mu, "A strategy of improving indoor air temperature prediction in HVAC system based on multivariate transfer entropy," *Building and Environment*, vol. 219, Jul. 2022, Art. no. 109164, https://doi.org/10.1016/j.buildenv.2022.109164.

[5] W. Zhang, Y. Wu, and J. K. Calautit, "A review on occupancy prediction through machine learning for enhancing energy efficiency, air quality and thermal comfort in the built environment," *Renewable and Sustainable Energy Reviews*, vol. 167, Oct. 2022, Art. no. 112704, https://doi.org/10.1016/j.rser.2022.112704.

[6] M. Marzouk and M. Atef, "Assessment of Indoor Air Quality in Academic Buildings Using IoT and Deep Learning," *Sustainability*, vol. 14, no. 12, Jan. 2022, Art. no. 7015, https://doi.org/10.3390/su14127015.

[7] A. Fonseca, I. Abreu, M. J. Guerreiro, C. Abreu, R. Silva, and N. Barros, "Indoor Air Quality and Sustainability Management—Case Study in Three Portuguese Healthcare Units," *Sustainability*, vol. 11, no. 1, Jan. 2019, Art. no. 101, https://doi.org/10.3390/su11010101.

[8] S. H. Ryu and H. J. Moon, "Development of an occupancy prediction model using indoor environmental data based on machine learning techniques," *Building and Environment*, vol. 107, pp. 1–9, Oct. 2016, https://doi.org/10.1016/j.buildenv.2016.06.039.

[9] C. Wang, D. Yan, H. Sun, and Y. Jiang, "A generalized probabilistic formula relating occupant behavior to environmental conditions," *Building and Environment*, vol. 95, pp. 53–62, Jan. 2016, https://doi.org/10.1016/j.buildenv.2015.09.004.

[10] C. Li, H. Zhu, X. Lian, Y. Liu, X. Li, and Y. Feng, "Study of 'time-lag' of occupant behavior occurrences for establishing an occupant-centric building control system," *Building and Environment*, vol. 216, May 2022, Art. no. 109005, https://doi.org/10.1016/j.buildenv.2022.109005.

[11] S.-Y. Lin, S.-C. Chiu, and W.-Y. Chen, "Simple automatic supervisory control system for office building based on energy-saving decoupling indoor comfort control," *Energy and Buildings*, vol. 86, pp. 7–15, Jan. 2015, https://doi.org/10.1016/j.enbuild.2014.09.066.

[12] S. Lee and D.-H. Choi, "Reinforcement Learning-Based Energy Management of Smart Home with Rooftop Solar Photovoltaic System, Energy Storage System, and Home Appliances," *Sensors*, vol. 19, no. 18, Jan. 2019, Art. no. 3937, https://doi.org/10.3390/s19183937.

[13] J. A. Weyn, D. R. Durran, and R. Caruana, "Can Machines Learn to Predict Weather? Using Deep Learning to Predict Gridded 500-hPa Geopotential Height From Historical Weather Data," *Journal of Advances in Modeling Earth Systems*, vol. 11, no. 8, pp. 2680–2693, 2019, https://doi.org/10.1029/2019MS001705.

[14] S. Serroni, M. Arnesano, L. Violini, and G. M. Revel, "An IoT measurement solution for continuous indoor environmental quality monitoring for buildings renovation," *Acta IMEKO*, vol. 10, no. 4, pp. 230–238, Dec. 2021, https://doi.org/10.21014/acta_imeko.v10i4.1182.

[15] T. Xing, K. Sun, and Q. Zhao, "MITP-Net: A deep-learning framework for short-term indoor temperature predictions in multi-zone buildings," *Building and Environment*, vol. 239, Jul. 2023, Art. no. 110388, https://doi.org/10.1016/j.buildenv.2023.110388.

[16] T. T. K. Tran, S. M. Bateni, S. J. Ki, and H. Vosoughifar, "A Review of Neural Networks for Air Temperature Forecasting," *Water*, vol. 13, no. 9, Jan. 2021, Art. no. 1294, https://doi.org/10.3390/w13091294.

[17] S. Francik and S. Kurpaska, "The Use of Artificial Neural Networks for Forecasting of Air Temperature inside a Heated Foil Tunnel," *Sensors*, vol. 20, no. 3, Jan. 2020, Art. no. 652, https://doi.org/10.3390/s20030652.

[18] S. R. Gopi and M. Karthikeyan, "Effectiveness of Crop Recommendation and Yield Prediction using Hybrid Moth Flame Optimization with Machine Learning," *Engineering, Technology & Applied Science Research*, vol. 13, no. 4, pp. 11360–11365, Aug. 2023, https://doi.org/10.48084/etasr.6092.

[19] A. H. Blasi, M. A. Abbadi, and R. Al-Huweimel, "Machine Learning Approach for an Automatic Irrigation System in Southern Jordan Valley," *Engineering, Technology & Applied Science Research*, vol. 11, no. 1, pp. 6609–6613, Feb. 2021, https://doi.org/10.48084/etasr.3944.

[20] P. V. D. de Souza, L. P. de Rezende, A. P. Duarte, and G. V. Miranda, "Maize Yield Prediction using Artificial Neural Networks based on a Trial Network Dataset," *Engineering, Technology & Applied Science Research*, vol. 13, no. 2, pp. 10338–10346, Apr. 2023, https://doi.org/10.48084/etasr.5664.

[21] E. Mocanu, P. H. Nguyen, M. Gibescu, and W. L. Kling, "Deep learning for estimating building energy consumption," *Sustainable Energy, Grids and Networks*, vol. 6, pp. 91–99, Jun. 2016, https://doi.org/10.1016/j.segan.2016.02.005.

[22] J. Cifuentes, G. Marulanda, A. Bello, and J. Reneses, "Air Temperature Forecasting Using Machine Learning Techniques: A Review," *Energies*, vol. 13, no. 16, Jan. 2020, Art. no. 4215, https://doi.org/10.3390/en13164215.

[23] A. M. Gómez-Orellana, D. Guijo-Rubio, J. Pérez-Aracil, P. A. Gutiérrez, S. Salcedo-Sanz, and C. Hervás-Martínez, "One month in advance prediction of air temperature from Reanalysis data with eXplainable Artificial Intelligence techniques," *Atmospheric Research*, vol. 284, Mar. 2023, Art. no. 106608, https://doi.org/10.1016/j.atmosres.2023.106608.

[24] T. Alshammari, R. A. Ramadan, and A. Ahmad, "Temporal Variations Dataset for Indoor Environmental Parameters in Northern Saudi Arabia," *Applied Sciences*, vol. 13, no. 12, Jan. 2023, Art. no. 7326, https://doi.org/10.3390/app13127326.

[25] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," *PeerJ. Computer Science*, vol. 7, 2021, Art. no. e623, https://doi.org/10.7717/peerj-cs.623.

[26] M. Abdurohman, A. G. Putrada, and M. M. Deris, "A Robust Internet of Things-Based Aquarium Control System Using Decision Tree Regression Algorithm," *IEEE Access*, vol. 10, pp. 56937–56951, 2022, https://doi.org/10.1109/ACCESS.2022.3177225.

[27] "scikit-learn: machine learning in Python — scikit-learn 1.4.1 documentation." [Online]. Available: https://scikit-learn.org/stable/.

[28] S. Murugan, B. M. Kumar, and S. Amudha, "Classification and Prediction of Breast Cancer using Linear Regression, Decision Tree and Random Forest," in *2017 International Conference on Current Trends in Computer, Electrical, Electronics and Communication (CTCEEC)*, Mysore, India, Sep. 2017, pp. 763–766, https://doi.org/10.1109/CTCEEC.2017.8455058.

[29] H. Salem, A. E. Kabeel, E. M. S. El-Said, and O. M. Elzeki, "Predictive modelling for solar power-driven hybrid desalination system using artificial neural network regression with Adam optimization," *Desalination*, vol. 522, Jan. 2022, Art. no. 115411, https://doi.org/10.1016/j.desal.2021.115411.

[30] T. Yan, S.-L. Shen, A. Zhou, and X. Chen, "Prediction of geological characteristics from shield operational parameters by integrating grid search and *K*-fold cross validation into stacking classification algorithm," *Journal of Rock Mechanics and Geotechnical Engineering*, vol. 14, no. 4, pp. 1292–1303, Aug. 2022, https://doi.org/10.1016/j.jrmge.2022.03.002.