ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

DEEP LEARNING TECHNIQUES TO ENHANCE ENERGY EFFICIENCY OF HOME APPLIANCES BY ANALYSING AIR QUALITY LEVELS

Jasbir Singh Saini[1], Dr. Sunny Arora[2], Dr. Sushil Kamboj[3]

[1]*Research Scholrar, Deptt. of CSE,Guru Kashi University,* Talwandi Sabo, Bathinda
sainijassbir@gmail.com

[2]*Professor, Deptt. of CSE Guru Kashi University* Talwandi Sabo, Bathinda
sunnyarora@gku.ac.in

[3]*Professor Deptt.of IT CGC, Landran* Chandigarh dr.sushilkamboj@gmail.com

ABSTACT:

Energy efficiency in home appliances is a critical area of research that addresses the growing demand for reducing energy consumption. The rapid growth in artificial intelligence has prioritized the development of advanced methods to improve sustainable energy consumption, particularly by optimizing the energy efficiency of home appliances. This paper introduces a novel deep learning-based framework to enhance energy efficiency in home appliances by leveraging insights from indoor air quality (IAQ) metrics. Unlike conventional energy management approaches that face challenges such as limited datasets, computational inefficiencies, and lack of generalizability, this research incorporates advanced preprocessing and augmentation techniques. Specifically, a hybrid SMOTE-ENN approach addresses class imbalance, while Z-score normalization ensures consistent feature scaling. Among the evaluated models, Bidirectional GRU and Stacked LSTM stand out, achieving exceptional validation accuracies of 99.81% and 99.64%, respectively, demonstrating superior generalization. This framework uniquely integrates indoor air quality data to optimize energy usage dynamically, showcasing how environmental factors such as CO2, humidity, and temperature can inform sustainable energy practices. These findings underscore the transformative potential of deep learning in fostering eco-friendly innovations for smart home energy management.

Keywords: *Home appliances, Energy efficiency, Deep Learning, Sustainable Living, SMOTE- ENN, Z-Score*

INTRODUCTION

Since the era of 21st century has started, the world has experienced a very strong growth in the field of global energy consumption in almost all the regions. It has been analysed that the consumption of energy often increases because of the factors like economic growth, expansion of demographic and higher usage of electricity per capita [1]. Optimizing the efficiency of household appliances plays a crucial role to mitigate this rise in energy demand because of its contribution of over 30% by the residential consumers in certain nations. It also addresses both the economic and

ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

environmental impacts of rising energy consumption which signifies the importance of sustainable practices in residential energy management [2]. Moreover, such improvements also contribute in the reduction of pollution along with the individual contributions to climate change. In fact, an energy-efficient home integrates some advanced technologies and designs which lessens the using up of energy as well as maintains the same level of safety, convenience, comfort, and visual attractiveness as traditional homes [3]. Figure 1 displays the percent of energy consumed by the home appliances.

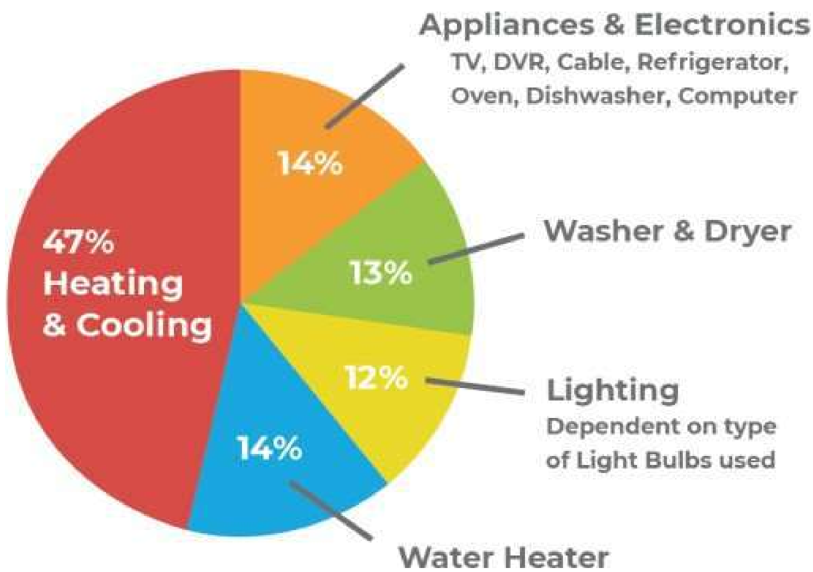


Figure 1: Energy contributions by different home appliances to highlight their share in total residential energy consumption [31]

Globally, policies aimed at enhancing energy efficiency provide significant benefits to both energy suppliers and consumers, yielding environmental, social, and economic advantages. Efficient energy use is essential for strengthening energy supply amid rising demand and economic expansion. This includes advancements in energy management to mitigate peak energy demands, optimize household appliance usage, and develop increasingly efficient appliances [4]. Particular importance is placed on improving the efficiency of heating and cooling systems, which constitute significant energy loads within residential settings. The Air Quality Index (AQI) also serves as an important metric to optimize the performance and energy efficiency of household appliances. Increased levels of pollutants like dust can affect the effectiveness of air purifier and fans which requires adjustments to maintain optimal indoor air quality. Similarly, higher CO2 levels, indicative of increased occupancy, may require optimal use of heating, ventilation, and lighting systems for efficient operation [5]. Analyzing past indoor air quality index data enables to identify the patterns and correlations, which facilitate the personalized recommendations to

ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

optimize appliance performance by considering air quality conditions. Thus, using an indoor air quality index will help us to predict how our appliances will perform and aid us to plan ahead to keep them working efficiently in different indoor conditions [6].

There are various conventional techniques to enhance energy efficiency of home appliances which involve implementation of energy-efficiency standards, improvement of appliance design, and promotion of energy-efficient technologies. However, these methods also face challenges in terms of complexity of appliance design, resistance to standards and regulations from manufacturers, and the influence of consumer behaviour [7]. To address this, deep learning provides a promising solution by optimizing appliance design through data-driven insights. Apart from this, deep learning algorithms can identify correlations and patterns by analyzing large datasets on appliance performance, energy consumption, and IAQ that traditional methods may overlook. This enables the development of more efficient appliance designs that not only save energy but also contribute to improved IAQ [8].

There are various studies which have shown their contribution in the energy efficiency of home appliances using machine and deep learning techniques. [9] Proposed an Artificial Intelligence-based Energy Management Model for green buildings in order to prioritize comfort and safety of user along with the energy efficiency. They used universal infrared communication system and LSTM model for optimizing the energy consumption and emphasized HVAC system airside design optimization to display economic and environmental benefits. Green buildings benefit from the AI-EMM computed high- performance ratio of 94.3%, reduced energy consumption of 15.7% , accuracy (97.4%), energy management level (95.7%), and prediction accuracy (97.1%). [10] used correlation analysis to collect the data to discard redundant sensors and focused to optimize IoT system design for smart homes. They used data analysis and prediction technique to enhance energy efficiency by correlating heterogeneous IoT sensor data and proposed a machine learning- based intelligent service model which was evaluated using RMSE. The results indicated that the gradient-boosting regressor was the most effective by achieving 22.29 as RMSE. Besides this, various architectures of deep recurrent neural networks (DRNNs) are explored those are tailored for medium- and long-term energy demand predictions, specifically for heating and electricity consumption at a 1-hour resolution. Their proposed DRNN model surpassed support vector machine (SVM) and gradient boosting (GB) regression models by 5.4% and 7.0%, respectively by showcasing its superior performance in energy forecasting accuracy. a novel model is proposed which consisted of three components i.e. smoothing which employed Kalman filter for eliminating the noise from data, optimization to minimize the cost error in real time data by using firefly and genetic algorithms, and control to manage distribution of energy for lightning, temperature etc. [12]. efficiently using Mamdani fuzzy logics. The researchers also compared their work with the existing techniques

ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

and found that their model outperformed as well as they highlighted the importance of using optimizers for energy efficiency and improving user comfort along with the impact of adaptive controllers to overcome incorrect PID controller selections. Shree, Lakshmi, et al. sourced the data from Kaggle which comprised of 29 features to focus on the minimal consumption of energy using various machine learning models such as LSTM along with the optimization techniques like genetic algorithm and grey wolf optimization to fine-tune hyper parameters [13]. While evaluating, GWO-LSTM highlighted the superior performance which showcased its exceptional predictive capabilities with minimal errors. Khan et al. used one dimensional deep convolutional neural network, LSTM, and scheduling algorithm to extract features, load forecasting based on features that had been extracted and optimize appliance operation times respectively to develop an energy consumption control system for smart homes [14]. They validated their model through simulation scenarios with authentic datasets which demonstrated its effectiveness in meeting energy demands without requiring additional energy sources and found that their proposed system displayed advancement in smart home energy management [14]. Liao et al. (2020) [32] explored deep learning techniques for air quality forecasting, focusing on CNNs, RNNs, LSTMs, and spatiotemporal networks for modeling nonlinear spatiotemporal features. They also discussed about the challenges like overfitting and practical implications for real-world deployment.

While existing research has made significant strides in developing models to enhance energy efficiency in smart rooms, several challenges and limitations persist. Firstly, there is often a lack of consideration for the substantial computational time required to train these models, which can impact practical implementation. Researchers also encounter challenges related to the availability of limited datasets, leading to potential biases or issues with generalizability. Furthermore, data inconsistency across different sources poses a significant challenge, affecting the reliability and accuracy of the models. Addressing these challenges will be crucial for advancing the field and developing more effective strategies for enhancing the energy efficiency of home appliances in smart rooms.

While existing studies have explored various methods to optimize energy efficiency in residential settings, the integration of indoor air quality (IAQ) as a critical factor remains underexplored. This research bridges that gap by incorporating IAQ parameters—such as CO2 levels, humidity, and temperature—into a novel deep learning framework, enabling dynamic and context-aware energy management. By employing advanced models like Bidirectional GRU, Stacked LSTM etc., this paper captures both short-term and long-term temporal dependencies, ensuring accurate predictions of appliance performance under varying environmental conditions. The use of a hybrid SMOTE-ENN technique to address class imbalance further enhances the reliability and generalizability of the results. This approach not only highlights the synergy between IAQ and energy efficiency but also sets a foundation for scalable and adaptive energy solutions in smart homes. Such work underscore the

ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

transformative potential of integrating IAQ metrics into sustainable energy management practices, paving the way for smarter, healthier, and more energy-efficient residential environments.

Contribution of the paper

In this paper the aim is to develop an automated system that uses deep learning techniques to identify as well as classify the air quality level based on multiple parameters including indoor air quality index. The contribution to perform the research is as follows:

- Initially, a dataset consisting of 132007 records with seven attributes, such as CO2 levels, humidity, PIR (Passive Infrared), temperature, indoor air quality indexes, and air quality level of rooms, were collected from two rooms 415 (Data I) and 776 (Data II).
- Subsequently, the data was preprocessed to check for null or missing values, followed by graphical visualization to understand the pattern of the dataset.
- To address the class imbalance issue, the SMOTE (Synthetic Minority Over-sampling Technique) technique was employed, and the features of the dataset were standardized through scaling.
- Various deep learning techniques were applied and trained with the dataset. The performances of these techniques were later examined using various standard metrics including the learning curves, confusion matrix, and computational time.

Structure of the paper

The paper is structured to first address energy efficiency and its societal implications along with the analysis of existing methodologies in Section I. Section II then delves into the methodology employed to develop an energy efficiency model using advanced learning approaches. Subsequently, Section III examines the performance of these classifiers. Finally, Section IV concludes the entire paper and offers a comprehensive summary of the findings and their broader implications.

I. RESEARCH METHOD

This section defines the phases that have been used to predict and classify the Air Quality Level of a room using hybrid advanced deep learning techniques, as shown in Figure 2.

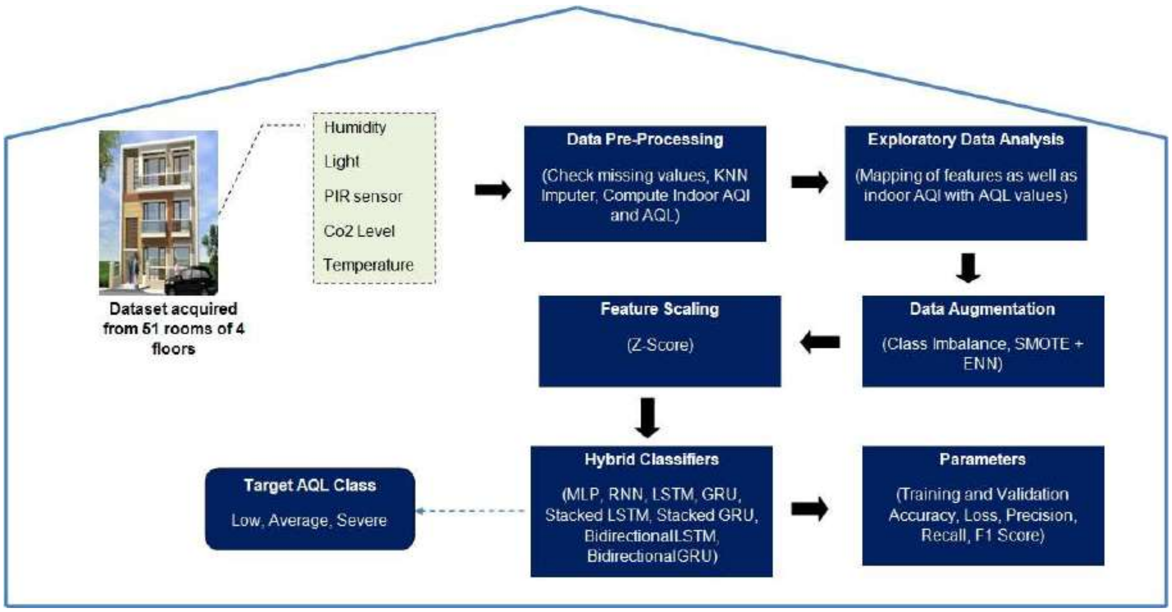


Figure 2: Proposed system for air quality assessment using deep learning classifiers

Dataset: The dataset collected records from 255 sensors which have been located across fifty one rooms spanning 4 floors of Sutardja Dai Hall at UC Berkeley. It includes diverse attributes like PIR sensor, Carbon Dioxide (CO2) tiers, humidity, temperature, and luminosity, with readings recorded every five seconds time series information in the form of UNIX EPOCH TIME timestamp [15].

Data Preprocessing: During this phase, initially, a thorough analysis is carried out to identify any missing or null values in each attribute for all 51 rooms. This is done to make sure that the data is complete and accurate, as shown in Table 1. The KNN imputer technique was employed for filling the missing values on the basis of information fetched from nearby data points [16]. This approach helps maintain the structure and patterns within the dataset. Afterward, using a comprehensive dataset, the air quality index (AQI) values for each room were calculated. This entailed collecting and analyzing the several characteristics linked to each space in order to calculate a comprehensive measure that represents the air quality.

However, it was noted throughout this procedure that certain estimated AQI values were negative, which are not suitable for meaningful interpretation and analysis. In order to tackle this problem, the dataset was improved by removing records with negative AQI values and concentrating only on the data from two rooms i.e. 415 (Data-I) and 776 (Data-II) that were chosen randomly from the original 51 rooms. Subsequently, utilizing the AQI data, we derived the AQL values and established our

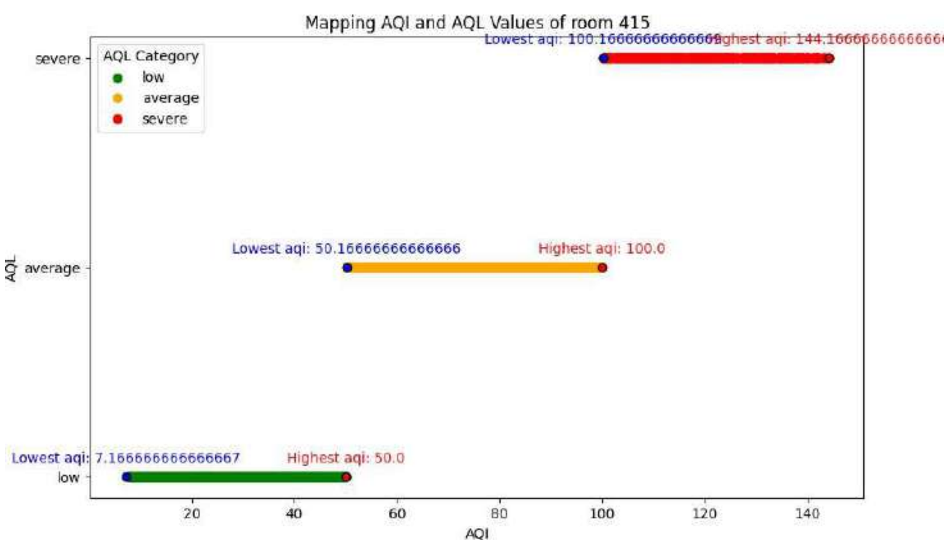
ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

desired classes as Low (0-50), Average (51-100), and Severe (101-500).

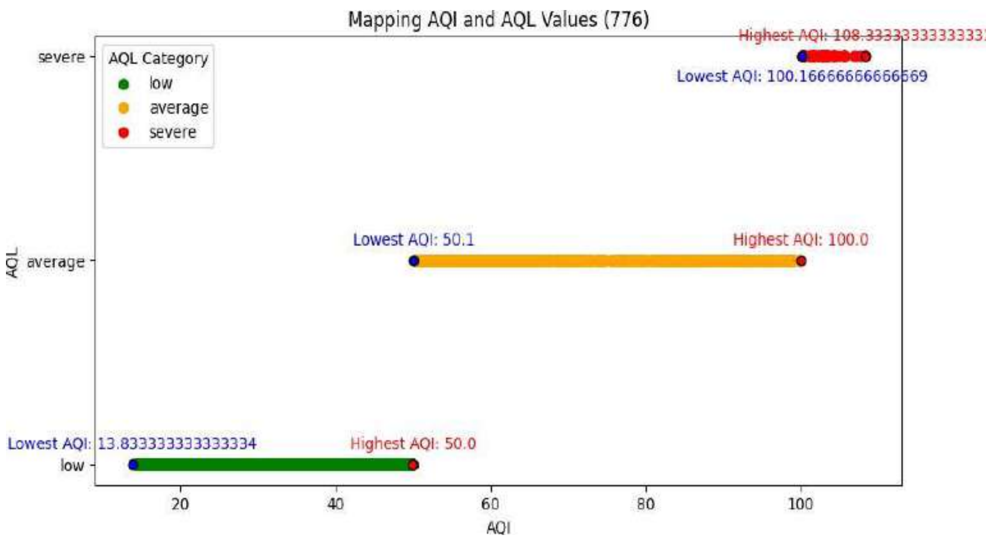
Exploratory Data Analysis: In this paper, EDA is used to uncover some crucial information in order to understand the complex relationships between energy consumption as well as various environmental factors. Figure 3 outlines the correlation between computed indoor air quality index (AQI) values and the corresponding air quality level classes, such as LOW, AVERAGE, and SEVERE. The purpose is to determine the minimum and maximum indoor AQI values recorded from both rooms 415 (Data I) and 776 (Data II). This information is crucial for improving the classification system and establishing distinct thresholds for different air quality levels. Ultimately, this will enhance comprehension and enable effective management of indoor environmental conditions.

Table 1: Comparison of missing values across attributes in Data-I and Data-II, to emphasize the significance of imputing missing data for accurate analysis

Attributes	Data-I (415)	Data-II (776)
Co2	0	1095
Humidity	1113	1
Light	1113	1
PIR	55875	59171
Temperature	1114	0



(a)

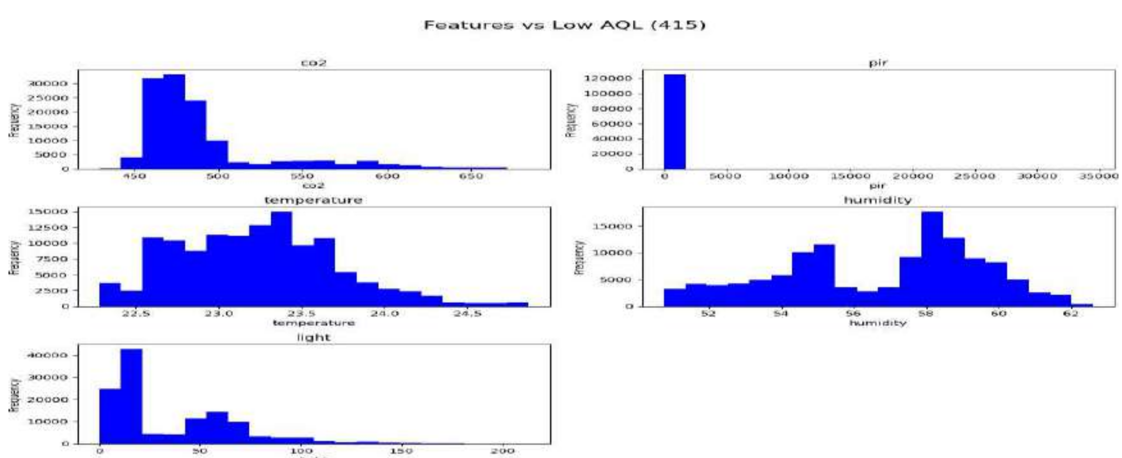


(b)

Figure 3: AQI values of Data I and Data II for different categories of AQL

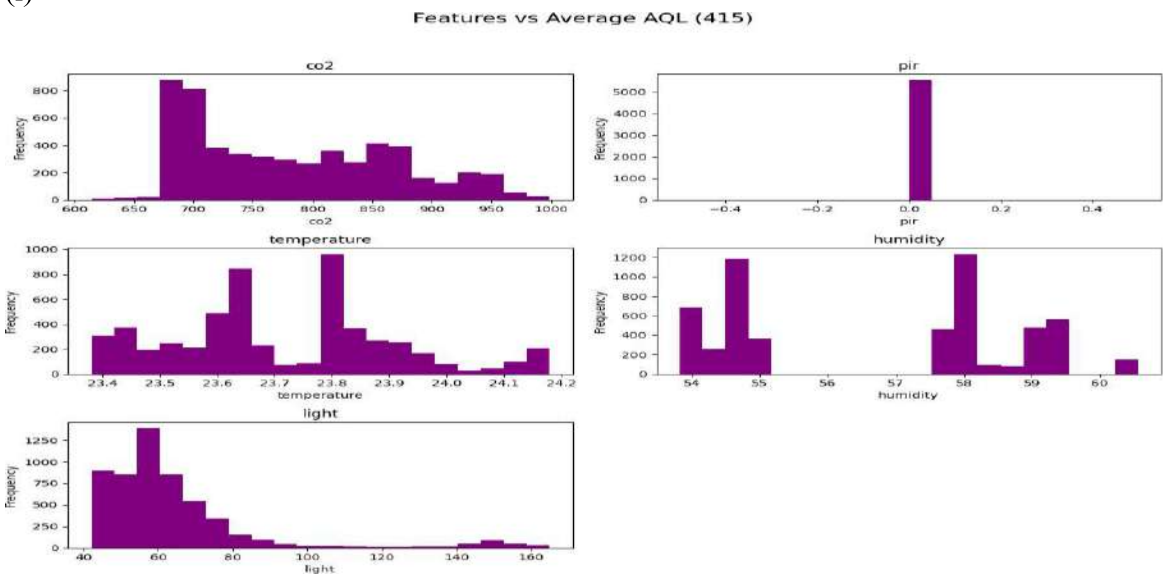
The graphical depiction in Figure 4 portrays the distribution of attribute values, encompassing CO2 concentration, PIR (Passive Infrared), light intensity, humidity, and temperature, across varying levels of air quality: low, average, and severe. Upon analyzing the data of Room 415 in Figure 4(a), it is discerned that the highest frequency of occurrences indicating low air quality is observed within the intervals of 450 to 500 for CO2 concentration, 23.0 to 23.5 for temperature, 58 to 60 for humidity, and 0 to 25 for light intensity.

Similarly, for the average air quality category, peak incidences manifest within the ranges of 690 to 700 for CO2, 23.5 to 23.9 for temperature, 54 to 55 and 58 to 58.5 for humidity, and 40 to 80 for light intensity. Conversely, in the severe air quality classification, predominant values are recorded between 1050 to 1100 for CO2 concentration, 23.65 to 23.70 for temperature, below 58 (e.g., 57.84) for humidity, and 40 to 50 for light intensity. It's noteworthy that these values are approximations and not fixed constants.

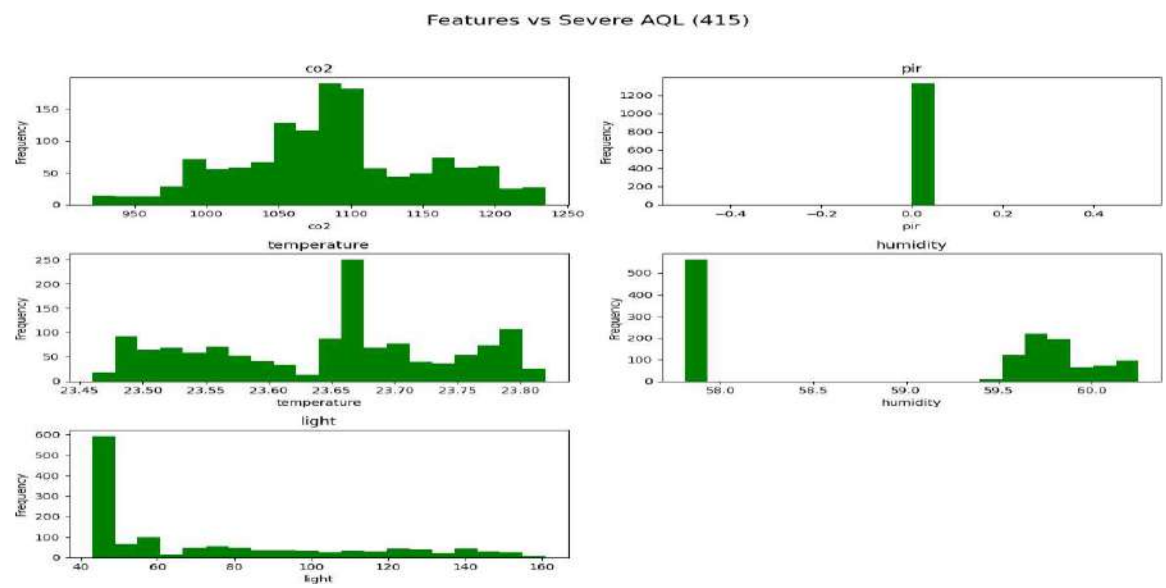


ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

(i)



(ii)



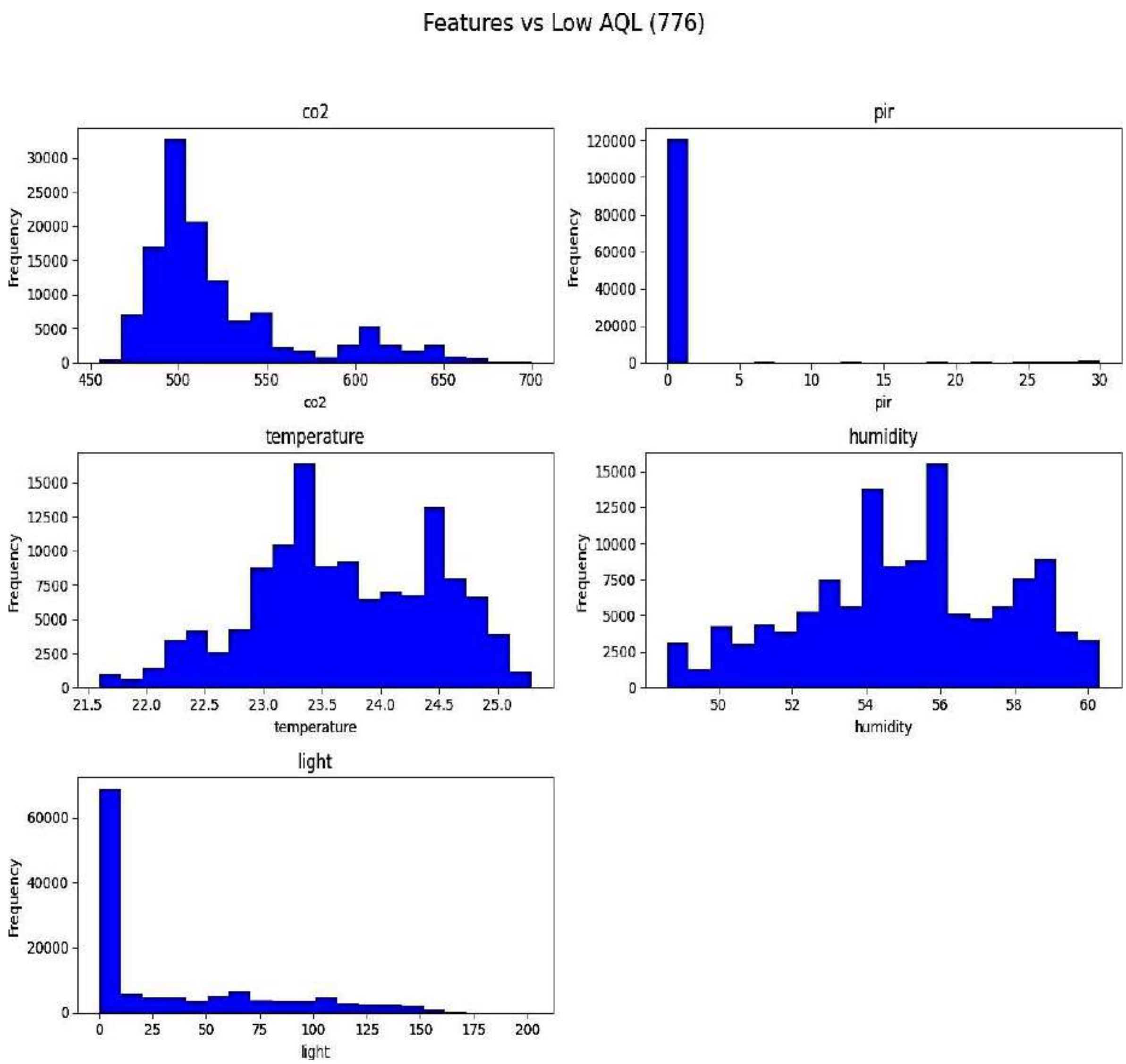
(iii)

Figure 4(a): Distribution of values of Data I features across various classes of AQL

Similarly, we have extracted analogous information from the data recorded for Room 776, as illustrated in Figure 4(b). The highest frequency of occurrences indicating low air quality is observed within the intervals of 450 to 500 for CO2 concentration,

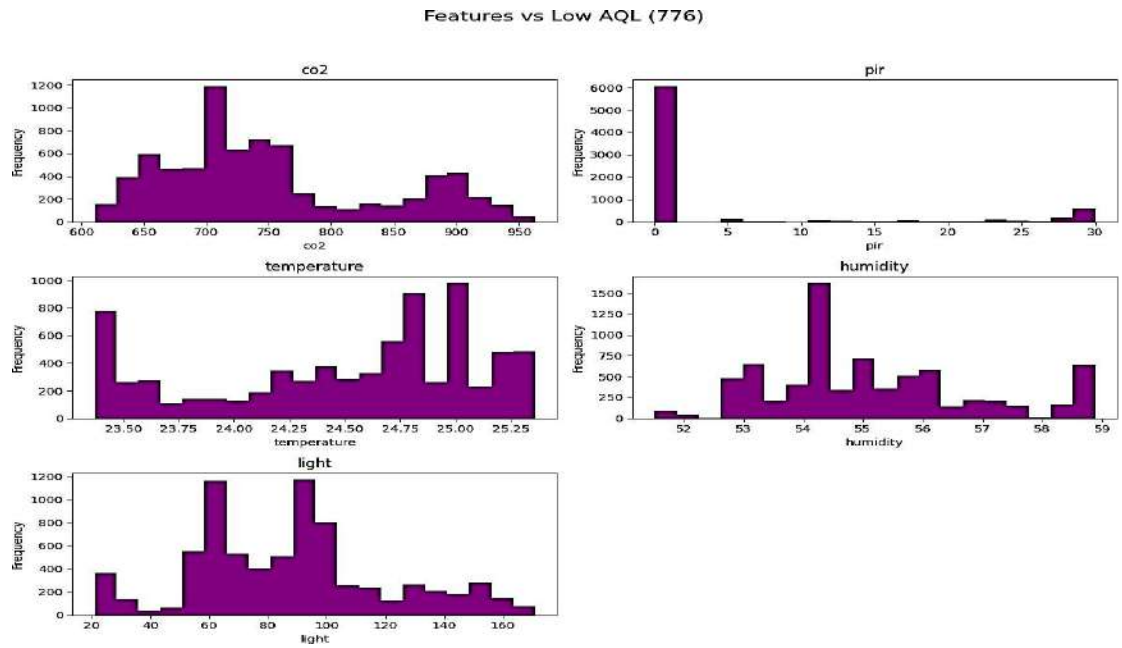
ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

23.0 to 23.5 for temperature, 57 to 58 for humidity, and 0 to 10 for light intensity. For the average air quality category, peak incidences manifest within the ranges of 700 to 720 for CO2, around 25.0 for temperature, 54 to 55 for humidity, and 60, and 90 to 100 for light intensity. Conversely, inthe severe air quality classification, predominant values are recorded at approximately 700 for CO2 concentration, around 25.2 for temperature, 54.5 for humidity, and 80 to 100 for light intensity. Additionally, for PIR (Passive Infrared), there are some occurrences of values apart from zero across all the attributes. It is important to note that these values are approximations and not fixed constants.

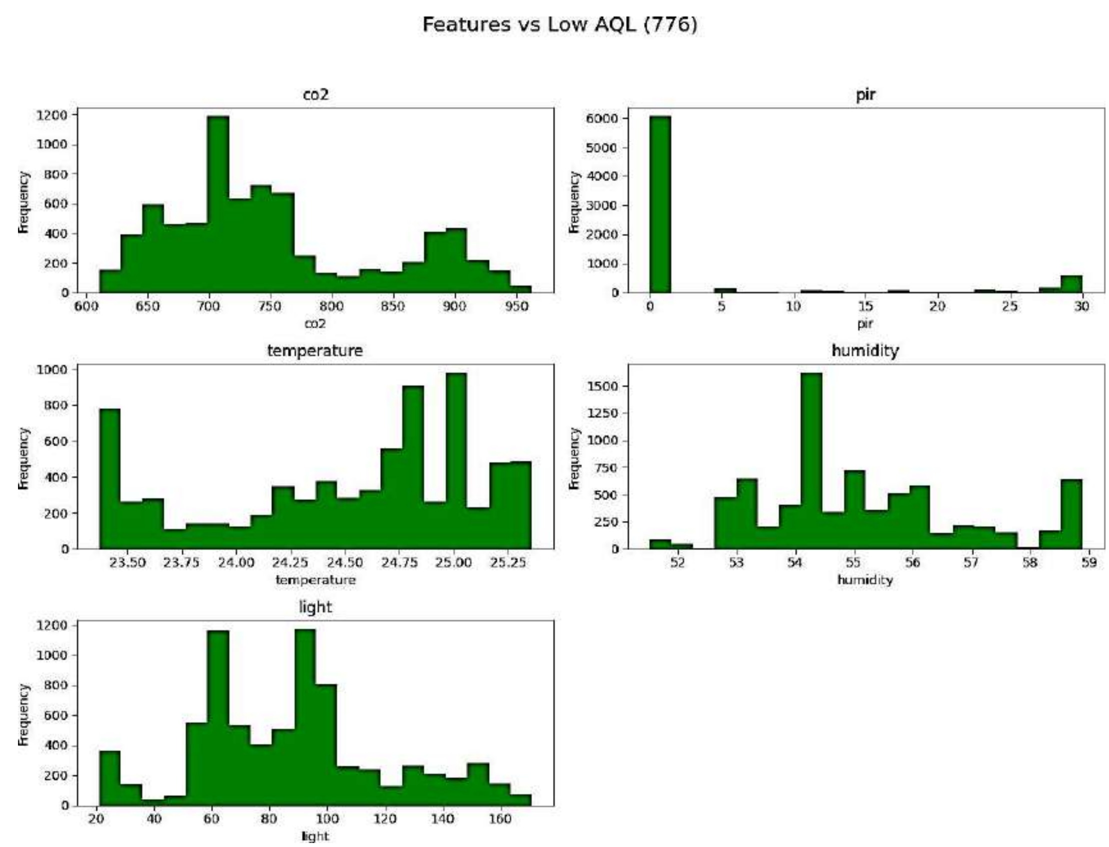


(i)

ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)



(ii)



(iii)

ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

Figure 4(b): Distribution of values of Data II features across various classes of AQLThe aim of these visualizations was to enhance energy efficiency, optimizing the detection and management systems based on these identified thresholds could facilitate proactive interventions, ensuring resource allocation aligns with actual environmental conditions, thereby minimizing energy consumption while maintaining air quality standards.

Data Augmentation: Hybrid approach of SMOTE and ENN has been used to augment the data by overcoming the issue of class imbalance, as shown in Figure 5. SMOTE generates synthetic instances for the minority class, which increases its representation in the dataset, while ENN removes noisy instances from both the minority and majority classes which leads to balanced and cleaner dataset overall [17]. By effectively addressing class imbalance and reducing noise, SMOTE-ENN enhances the performance of applied learning models. It can be represented as (i)

$$SMOTE + ENN\ (X, y) = ENN(SMOTE(X_{minority}), y_{minority}, k) \quad \text{(i)}$$

Here, X refers to feature matrix of the dataset, y implies target vector of class labels, and k is the number of nearest neighbours used in both SMOTE and ENN. However, many standard oversampling techniques, like standalone SMOTE, are effective in increasing the representation of minority classes by generating synthetic samples. But these methods can inadvertently introduce noise by creating synthetic samples near outliers or overlapping class boundaries, which can degrade model performance. Traditional under sampling methods, on the other hand, focus on removing data from the majority class to achieve balance, but this often results in loss of valuable information.

ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

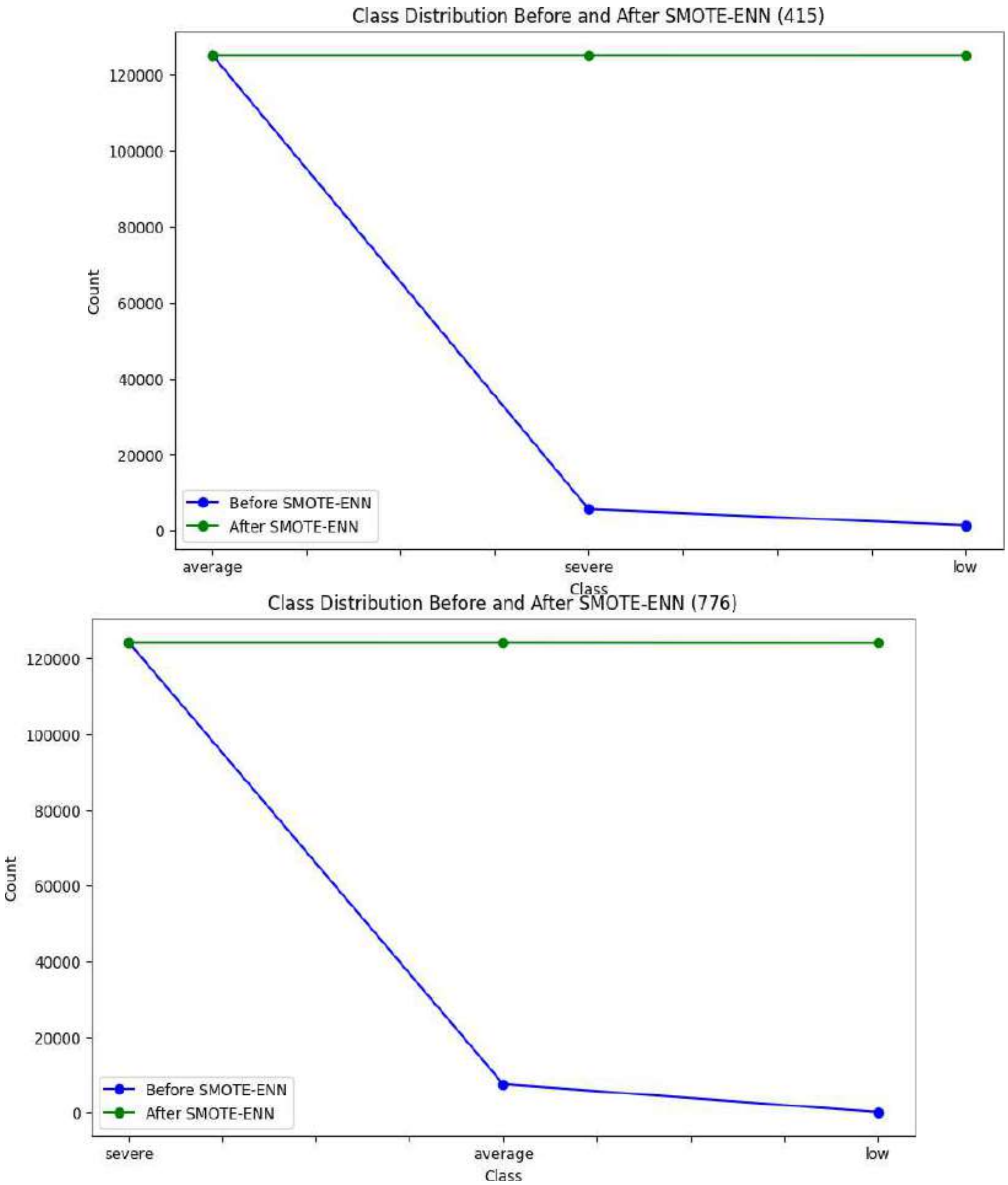


Figure 5: Class balancing achieved using the SMOTE-ENN technique to improve distribution of air quality levels in Data-I and Data-II

Feature Scaling: Z-scores are useful to identify outliers within a dataset. Data points with Z- scores significantly greater than or less than zero are considered outliers, as they deviate substantially from the average values of the dataset. By standardizing the data using Z-scores, it becomes easier to identify and understand the significance of outliers and to compare data points across different datasets with varying means and standard deviations [18], as shown in eq(ii).

ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

$$z = \frac{x - \mu}{\sigma} \quad \text{(ii)}$$

Here, x is the value of input data point, μ is the mean of population, and σ is the standard deviation of population.

Classifiers: In the context of indoor energy efficiency, the **MLP** structure is customized to analyze various environmental parameters, inclusive of temperature, humidity, occupancy, and lights situations, collected from sensors deployed within a building. The structure commonly includes an input layer, in which environmental data is fed into the model, followed with the aid of one or extra hidden layers, which perform nonlinear adjustments and feature extraction. Each neuron within those hidden layers applies weighted connections and activation capabilities to method the input records. Finally, an output layer produces predictions for energy efficient operation based totally at the discovered patterns [19]. The mathematical equation of MLP is represented as eq (iii)-(iv)

$$z_j^{(l)} = \sum_{i=1}^{n} w_{ij}^{l} x_i + b_j^{l} \quad \text{(iii)}$$

$$a_j^{(l)} = f(z_j^{(l)}) \quad \text{(iv)}$$

Here, n represents the input features, w_{ij}^{l} and $z_j^{(l)}$ refers to the weight connected to jth neuron, b_j^{l} is bias, x_i is the input feature, $f(.)$ is the activation function, and $a_j^{(l)}$ output of the jth neuron. **Recurrent Neural Networks** (RNNs) are a type of artificial neural network which are mostly used for processing sequential data to make them applicable for multiple tasks in optimizing energy efficiency for indoor environments. Its architecture includes input, hidden, and output layer with recurrent connections which enables the network for maintaining the information about past entered data in memory during processing current data [20]. By using the sequential nature of environmental data obtained from indoors such type of network contribute to optimize the use of energy by improvising the comfort of occupant as well as promote sustainability in indoor spaces. The hidden state (h_t) in RNN is defined by eq(v)

ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

$$h_t = \sigma\,(W_{hx}x_t + W_{hh}\,h_{t-1} + b_h) \quad \text{(v)}$$

Here, W_{hx} and W_{hh} implies the weight matrix for input to hidden and hidden to hidden connections, σ activation function, and b_h is the bias vector. **Long Short-Term Memory (LSTM)** networks is one variant of RNN architecture that's designed for taking long time dependencies in sequential primarily based information and addressing the problems related to vanished gradient. This property of LSTM makes it particularly efficient for the ones obligations which are associated with energy efficiency in indoor environments. The structure of LSTM includes memory cells having self-connected devices referred to as gates. These gates consist of an input, forget, and an output gate which alter the flow of data through the network and manipulate it at different levels of processing [21]. In the area of efficient use of indoor energy, LSTM networks excel at predicting complicated temporal styles in sensor data, which include fluctuations inside the temperature, tendencies of occupancy, and energy consumption profiles. The equations of the gates are presented as (vi-xi):

$$Forget\ gate\ (f_t) = \sigma(W_f.[h_{t-1}, x_t] + b_f) \quad \text{(vi)}$$

$$Input\ gate\ (i_t) = \sigma(W_i.[h_{t-1}, x_t] + b_i) \quad \text{(vii)}$$

$$Output\ gate\ (o_t) = \sigma(W_o.[h_{t-1}, x_t] + b_o) \quad \text{(viii)}$$

$$Candidate\ cell\ state\ (\tilde{c}_t) = tanh(W_c.[h_{t-1}, x_t] + b_c) \quad \text{(ix)}$$

$$Cell\ state\ update(c_t) = f_t\ .\ c_{t-1} + i_t.\tilde{c}_t \quad \text{(x)}$$

$$Hidden\ state\ update\ (h_t) = o_t\ .\ \tanh(c_t) \quad \text{(xi)}$$

Here, W_f, W_i, W_o, W_c are the weight matrices for the forget, input, output gate and candidate cell state respectively, σ and $tanh$ refers to the activation and the hyperbolic tangent activation function, (.) matrix multiplication. The other variation of recurrent neural network

ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

is **Gated Recurrent Unit structure**. This architecture is similar to LSTM however it copes with the restrictions of conventional RNNs. Apart from this; it is also able to efficiently fetch long-term dependencies in sequential facts. The structure of GRU is based totally on gating mechanisms which consist of reset gate and forget gate. These gates control the flow of records thru the network and allow the GRU for retaining or forget about data selectively from previous time steps. This characteristic permits the network to capture long-time period dependencies whilst reducing the vanishing gradient problem [22]. Historical records may be used to train GRUs network for learning the underlying patterns in addition to dynamics of indoor environments, which enable adaptive energy management techniques that work for conditions in real-time. It is mathematically represented by eqs (xii-xv):

$$Update\ gate\ (z_t) = \sigma(W_Z . [h_{t-1}, x_t] + b_Z) \quad \text{(xii)}$$

$$Reset\ gate\ (r_t) = \sigma(W_r . [h_{t-1}, x_t] + b_r) \quad \text{(xiii)}$$

$$Candidate\ hidden\ state\ (\tilde{h}_t) = tanh(W_h . [r_t \odot h_{t-1}, x_t] + b_h) \quad \text{(xiv)}$$

$$Hidden\ state\ update\ (h_t) = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad \text{(xv)}$$

Here, W_Z, W_r, W_h refers to the weight matrices of update, reset gate and candidate hidden state, $\odot$ implies to element wise multiplication, and $\tilde{h}_t$ is the current hidden state.

The architecture of **Bidirectional Long Short-Term Memory** (Bi-LSTM) networks incorporates records from both past and future time steps. It consists of two LSTM layers in which one layer approaches the input information in forward order and the other layer manners it in reverse order. While processing the data in both instructions, the Bi-LSTM captures data from both past and future contexts. This property of BiLSTM allows it to learn deeper representations of the enter records and enables it to higher recognize the temporal dynamics of indoor environmental data. Likewise, the architecture of a **Bi-GRU** consists of two GRU layers which fit precisely like BiLSTM. Here also, the bidirectional nature of the model allows it to fetch information from each beyond and future contexts concurrently to learn complex representations of the enter data [23,24]. The mathematical representation of

BiLSTM is shown in the form of forward (xvi) as well as backward direction (eq xvii):

Forward LSTM xvi

$$Forget\ gate\ (f_t^{(f)}) = \sigma(W_f^{(f)}.[h_{t-1}^{(f)}, x_t] + b_f^{(f)})$$

$$Input\ gate\ (i_t^{(f)}) = \sigma(W_i^{(f)}.[h_{t-1}^{(f)}, x_t] + b_i^{(f)})$$

$$Output\ gate\ (o_t^{(f)}) = \sigma(W_o^{(f)}.[h_{t-1}^{(f)}, x_t] + b_o^{(f)})$$

$$Candidate\ cell\ state\ (\tilde{c}_t^{(f)}) = tanh(W_c^{(f)}.[h_{t-1}^{(f)}, x_t] + b_c^{(f)})$$

$$Cell\ state\ update(c_t^{(f)}) = f_t^{(f)}.c_{t-1}^{(f)} + i_t^{(f)}.\tilde{c}_t^{(f)}$$

$$Hidden\ state\ update\ (h_t^{(f)}) = o_t^{(f)}.\tanh(c_t^{(f)})$$

xvii

Backward LSTM

$$Forget\ gate\ (f_t^{(b)}) = \sigma(W_f^{(b)}.[h_{t-1}^{(b)}, x_t] + b_f^{(b)})$$

$$Input\ gate\ (i_t^{(b)}) = \sigma(W_i^{(b)}.[h_{t-1}^{(b)}, x_t] + b_i^{(b)})$$

$$Output\ gate\ (o_t^{(b)}) = \sigma(W_o^{(b)}.[h_{t-1}^{(b)}, x_t] + b_o^{(b)})$$

$$Candidate\ cell\ state\ (\tilde{c}_t^{(b)}) = tanh(W_c^{(b)}.[h_{t-1}^{(b)}, x_t] + b_c^{(b)})$$

$$Cell\ state\ update(c_t^{(b)}) = f_t^{(b)}.c_{t-1}^{(b)} + i_t^{(b)}.\tilde{c}_t^{(b)}$$

$$Hidden\ state\ update\ (h_t^{(b)}) = o_t^{(b)}.\tanh(c_t^{(b)})$$

ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

Likewise, for Bi-GRU, the equations are represented as (xviii-xx)

Forward GRU xviii

$$Update\ gate\ (z_t^f) = \sigma(W_z^f.[h_{t-1}^f, x_t] + b_z^f)$$

$$Reset\ gate\ (r_t^f) = \sigma(W_r^f.[h_{t-1}^f, x_t] + b_r^f)$$

$$Candidate\ hidden\ state\ (\tilde{h}_t^f) = tanh(W_h^f.*r_t^f \odot h_{t-1}^f, x_t + + b_h^f)$$

$$Hidden\ state\ update\ (h_t^f) = (1 - z_t^f) \odot h_{t-1}^f + z_t^f \odot \tilde{h}_t^f$$

Backward GRU xix

$$Update\ gate\ (z_t^b) = \sigma(W_z^b.[h_{t-1}^b, x_t] + b_z^b)$$

$$Reset\ gate\ (r_t^b) = \sigma(W_r^b.[h_{t-1}^b, x_t] + b_r^b)$$

$$Candidate\ hidden\ state\ (\tilde{h}_t^b) = tanh(W_h^b.*r_t^b \odot h_{t-1}^b, x_t + + b_h^b)$$

$$Hidden\ state\ update\ (h_t^b) = (1 - z_t^b) \odot h_{t-1}^b + z_t^b \odot \tilde{h}_t^b$$

At the end, the output of both the networks at time stamp t are computed using the equation (xx)

Xx

$$h_t = [h_t^{(f)}, h_t^{(b)}]$$

Stacked Long Short-Term Memory (LSTM) architecture consists of multiple LSTM layers which are stacked on top of each other. Here, each LSTM layer processes the input data sequentially, where the output of one layer is served as the input to the next layer. When multiple LSTM layers are stacked, it allows the model to capture both short-term as well as long-term dependencies in the data to make it more capable of capturing the complex dynamics of indoor environmental variables. Likewise, **Stacked Gated Recurrent Unit networks** works on the same concept as stacked LSTM networks however uses GRU unitsrather. In this structure, various

ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

GRU layers are stacked on peak of every different to create a deep network. Every layer in a stacked GRU architecture includes a series of GRU units, with every unit having its own set of parameters to analyze the patterns and relationships of the input data. The output of one GRU layer serves the input to the following layer in order to allow network to learn hierarchical representations of the information across multiple stages of abstraction [25, 26]. In addition to this, the hyperparameter values used for training the models are also mentioned in Table 2.

Performance Metrics: In the context of energy efficiency for smart home appliances based on indoor air quality index and air quality level, various key metrics are typically used for evaluating the performance of classification models. Accuracy (xxi) provides an overall measure of model performance by taking the proportion of correctly classified instances [27].

Table 2: Hyperparameters and their selected values for model training

Hyperparameter	Value
Learning Rate	0.001
Batch Size	16
Epochs	15
Optimizer	Adam
Dropout Rate	0.5
Activation Layer	ReLU in hidden, Softmax in output

Loss (xxii) quantifies the difference between predicted and actual values. It is often measured using metrics like cross-entropy which serves as a gauge of model optimization and convergence [28]. In addition to these metrics, there are other measures also such as Precision
(xxiii) which is crucial to assess the ability of the model to avoid false positives to identify the equipment which is responsible for poor air quality. It is a measure of the proportion of correctly predicted positive cases among all predicted positive cases while as Recall (xxiv) is computing the ability of the model which captures all actual positive cases that are correctly identified and to balance the performance of the models on the basis of these metrics, F1 score (xxv) comes into play which is mostly useful when there is no synchronization between positive and negative instances [29, 30].

$$Acc = \frac{True\ Positive + True\ Negative}{True\ Positive + True\ Negative + False\ Positive + False\ Negative} \quad \text{xxi}$$

ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

$$Loss = \frac{(Actual\ Value - Predicted\ Value)2}{Number\ of\ observations} \quad \text{xxii}$$

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad \text{xxiii}$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad \text{xxiv}$$

$$F1\ score = 2\ \frac{Precision * Recall}{Recall + Precision} \quad \text{Xxv}$$

II. RESULTS AND DISCUSSION

The section presents the results of the models which have been trained by the data collected from both the rooms in different subsections.

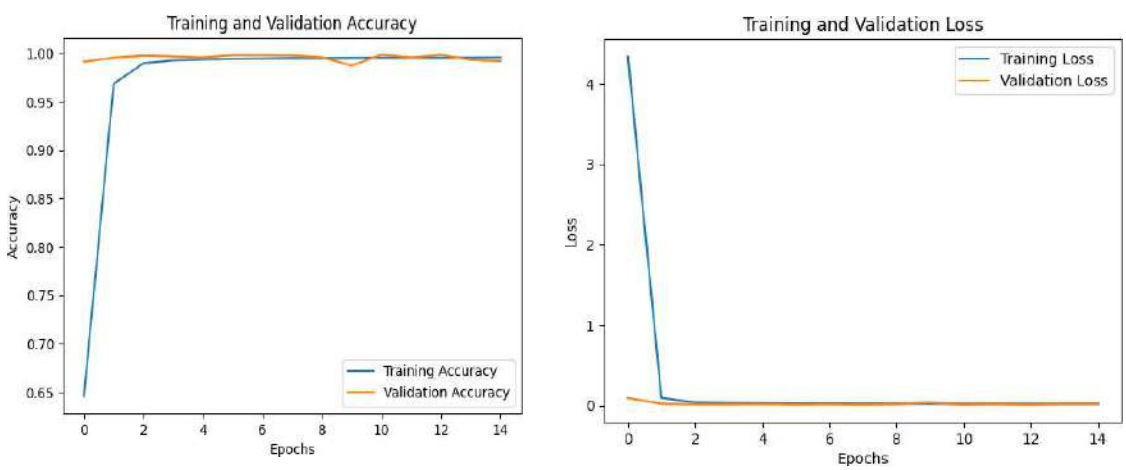
Analysis of models for the Data I of Room 415

Table 3 presents the performance metrics of various neural network models, including MLP (Multi-Layer Perceptron), RNN (Recurrent Neural Network), GRU (Gated Recurrent Unit), LSTM (Long Short-Term Memory), and combinations of these models, evaluated on training and validation datasets. The metrics considered are accuracy and loss, where higher accuracy and lower loss values indicate better performance.

Table 3: Accuracy and loss metrics for training and validation phases of various models applied to Data-I

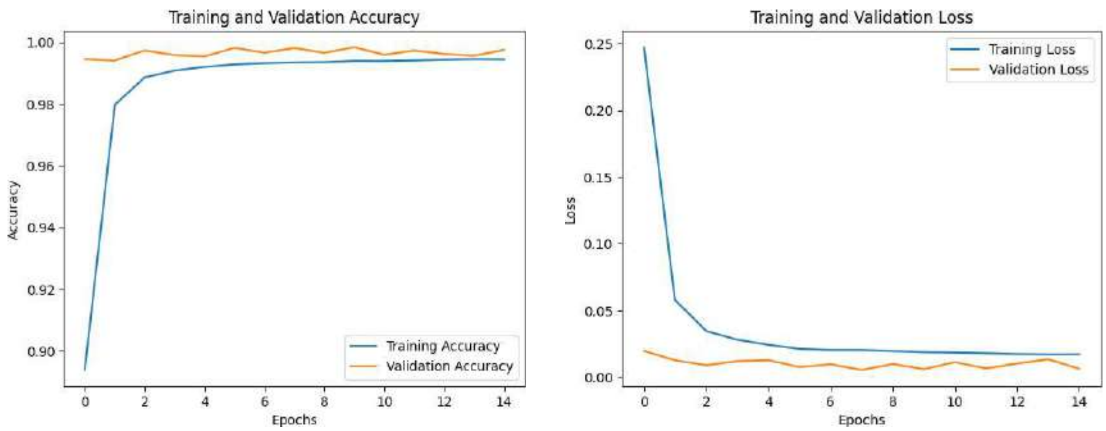
Models	Training		Validation	
	Acc	Loss	Acc	Loss
MLP	99.53	0.0146	99.16	0.0200
RNN	99.44	0.0169	99.75	0.0061
Bidirectional GRU	99.46	0.0152	99.81	0.0081
Bidirectional LSTM	99.47	0.0148	99.72	0.0066
LSTM	99.41	0.0166	99.57	0.0104
GRU	99.49	0.0149	99.79	0.0082
Stacked LSTM	99.40	0.0173	99.56	0.0102
Stacked GRU	99.45	0.0162	99.51	0.0156

Notably, in case of ***training phase***, the MLP model achieves the best accuracy at 99.53% with the lowest loss of 0.0146, suggesting it as the best performer in terms of training metrics as compared to the other models. Bidirectional versions of GRU and LSTM additionally show strong overall performance, with Bidirectional LSTM slightly outperforming Bidirectional GRU in terms of lowest loss (0.0148 vs. 0.0152) and marginally better accuracy (99.47% vs. 99.46%). This shows that the bidirectional nature of those models efficaciously captures temporal dependencies in both directions thereby enhancing their learning capability. Stacked LSTM and Stacked GRU architectures additionally carry out properly with 99.40% and 99.45% accuracies respectively, but slightly beneath their non-stacked architectures i.e. LSTM and GRU with 99.41% and 99.49% respectively, indicating that deeper networks do no longer necessarily enhance performance for this dataset. In case of ***validation phase***, The Bidirectional GRU and Bidirectional LSTM models achieve exceptionally high validation accuracy (99.81% and 99.72%, respectively) with low loss values (0.0081 and 0.0066, respectively), indicating strong generalization capabilities due to their ability to capture temporal dependencies in both directions. Models like the MLP achieve good performance with a validation accuracy of 99.16% and a loss of 0.0200, though they are slightly outperformed by RNN-based models due to their limited ability to capture sequential dependencies. LSTM and GRU models perform well, with accuracies of 99.57% and 99.79% and losses of 0.0104 and 0.0082, respectively, underscoring their effectiveness for sequential data. Stacked LSTM and Stacked GRU architectures exhibit solid performance but do not significantly surpass their single-layer counterparts, with accuracies of 99.56% and 99.51%, respectively.Figure 6 defines the learning curves of the models for training and validation accuracy as well as loss for 15 epochs. It has been found that MLP and RNN shows the good fit of learning curves while as remaining models shows some kind of zig-zag movements which indicates minor fluctuations in their performance. In addition to this, we can find here that the curve of validation loss is lower than the training loss which indicates that the validation dataset may be easier for the model to predict than the training dataset. Likewise, the curve of validation accuracy is higher than training accuracy from the beginning of the epochs and it implies that validation dataset may be easier or have a different distribution than the training set which leads to its higher performance.

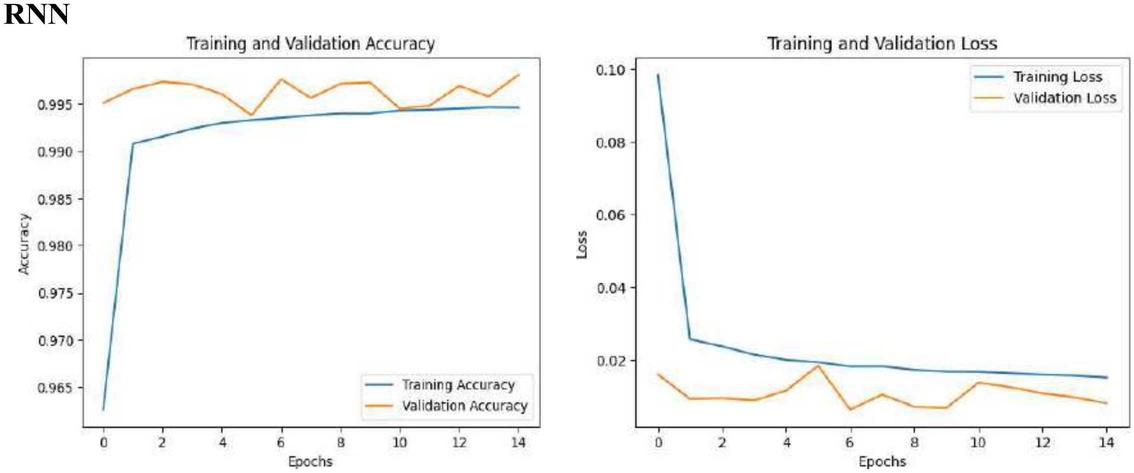


MLP

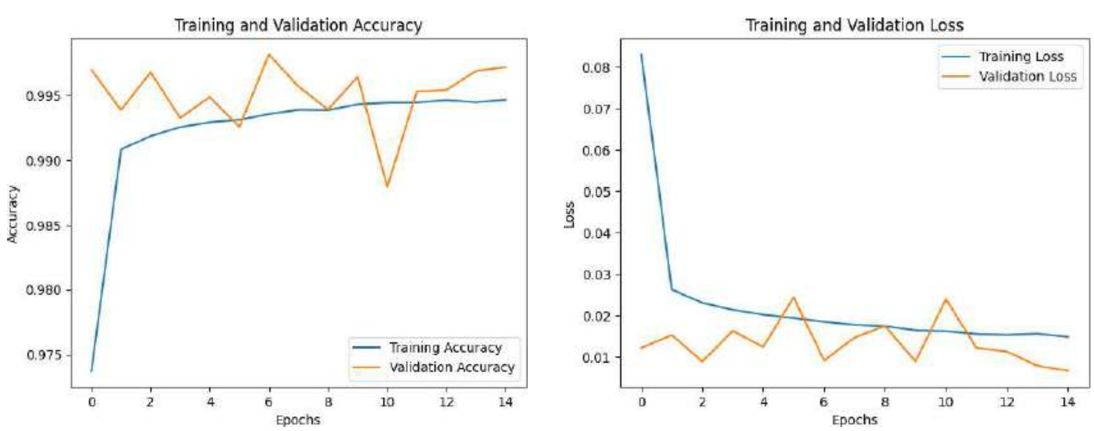
ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)



RNN

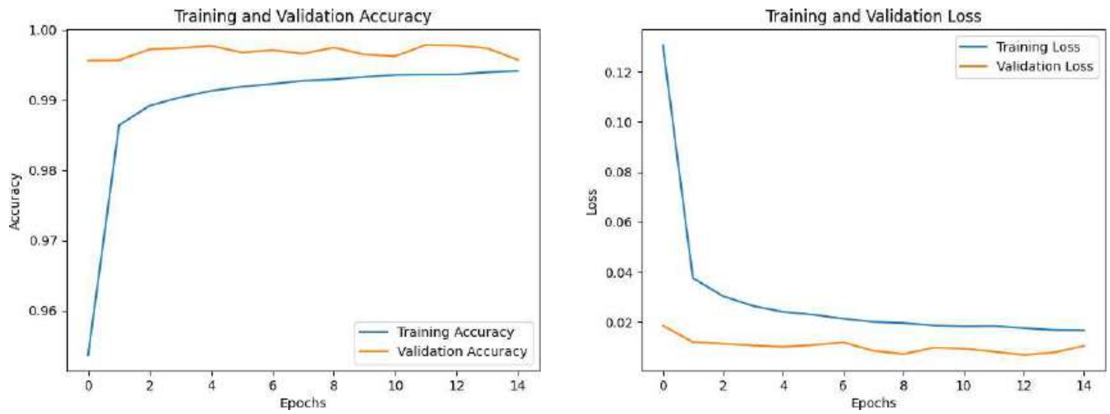


Bidirectional GRU

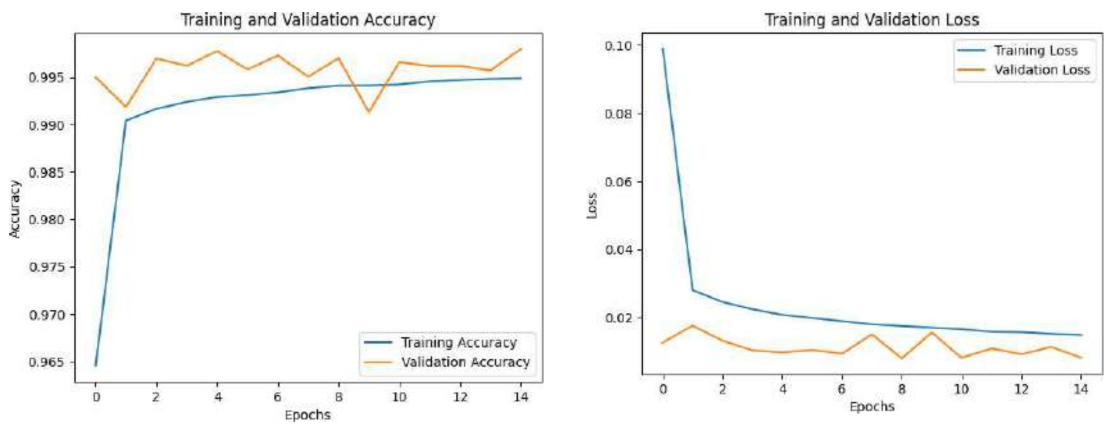


Bidirectional LSTM

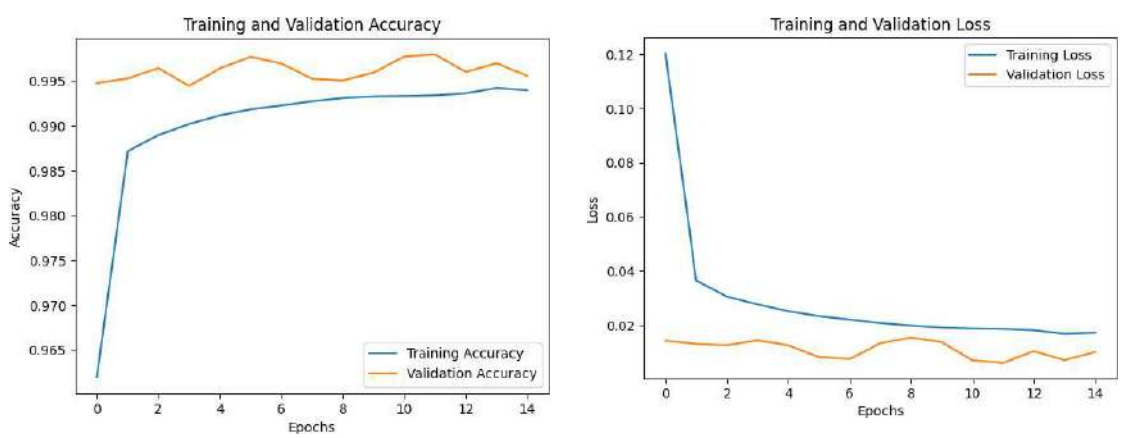
ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)



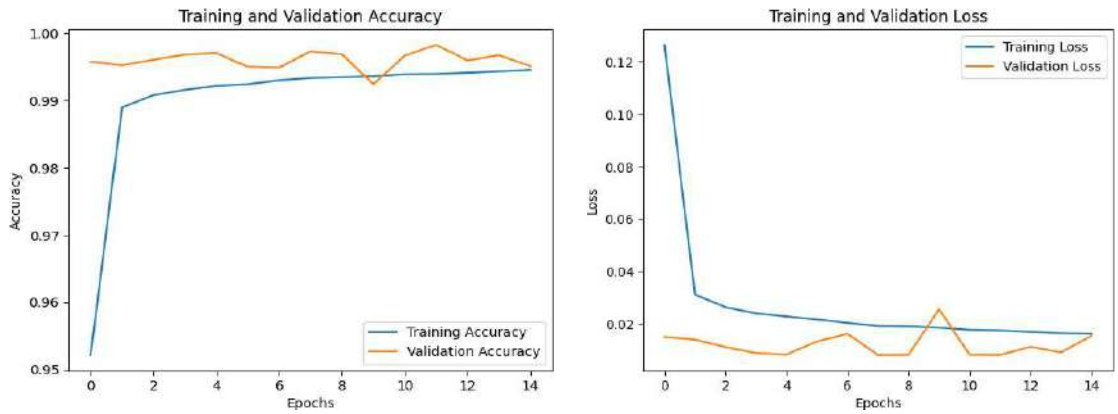
LSTM



GRU



Stacked LSTM



Stacked GRU

Figure 6: Learning curves depicting training/validation accuracy and loss for models applied to Data-I to indicate convergence and generalization trends.Table 4 presents precision, recall, and F1 scores for various neural network models, providing a comprehensive evaluation of their performance.

Table 4: Analysis of models for the energy efficient home appliances Data-I

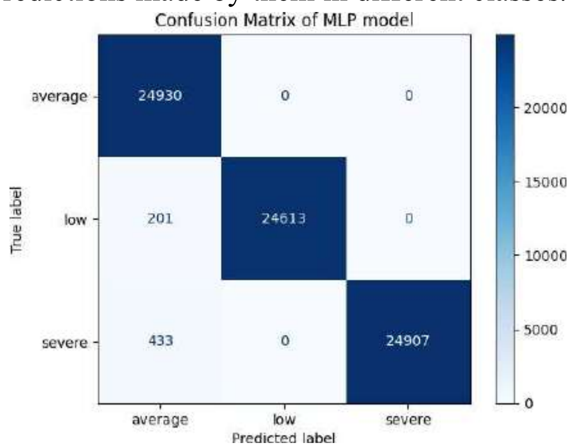
Models	**Precision**	**Recall**	**F1score**
MLP	0.9971	0.9915	0.9914
RNN	0.9975	0.9975	0.9974
Bidirectional GRU	0.9980	0.9982	0.9979
Bidirectional LSTM	0.9972	0.9972	0.9972
LSTM	0.9957	0.9957	0.9956
GRU	0.9979	0.9980	0.9979
Stacked LSTM	0.9955	0.9955	0.9954
Stacked GRU	0.9951	0.9950	0.9950

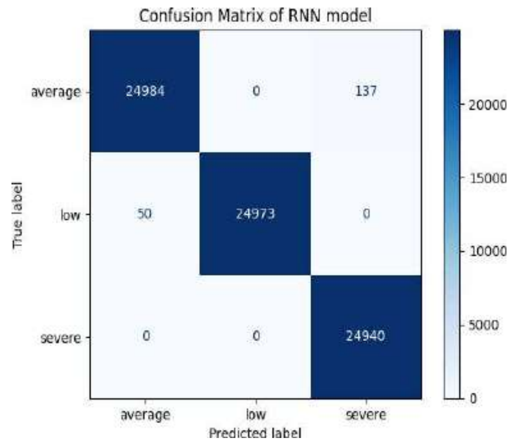
On comparing the performance of all the applied classifiers, it has been found that BidirectionalGRU computed the highest scores of precision, recall, and F1 score with 0.9980, 0.9982, and 0.9979 respectively followed by Gated Recurrent Unit with precision as 0.9979, recall as 0.9980, and F1 score as 0.9979. This performance depicts that these models have been able to classify the instances correctly. RNN, LSTM and BidirectionalLSTM also indicates their effectiveness by generating the balance values of precision (0.9975, 0.9957 and 0.9972), (0.9975, 0.9957, and 0.9972), and (0.9974, 0.9956, and 0.9972) respectively. Likewise, stacked LSTM and Stacked GRU also performed in a similar way by managing balance between precision, recall, and F1 score with their generated values. However, it has been seen that the least value has been generated by MLP in terms of recall and F1 score with 0.9915 and 0.9914 which shows that the model still needs to improve its performance

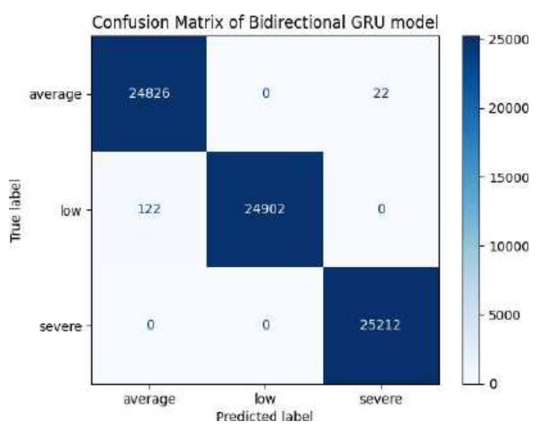
ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

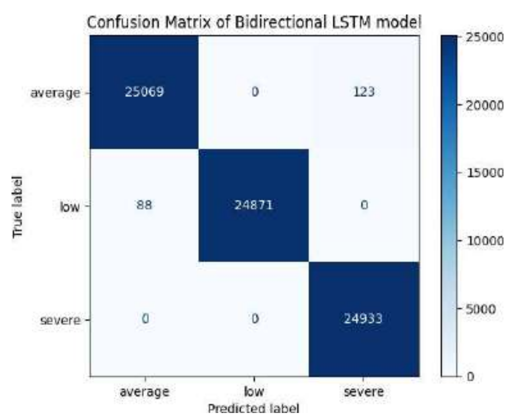
for this specific dataset. This indicates that the MLP requires further improvement to enhance its performance for this specific dataset.

A confusion matrix of 3x3 size has been also created in Figure 7 to analyse the actual and predicted values of the models on the basis of their true positive, true negative, false positive, and false negative values. Additionally, it also enables us to understand the strengths and the weakness of the models by separating the predictions made by them in different classes.









ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

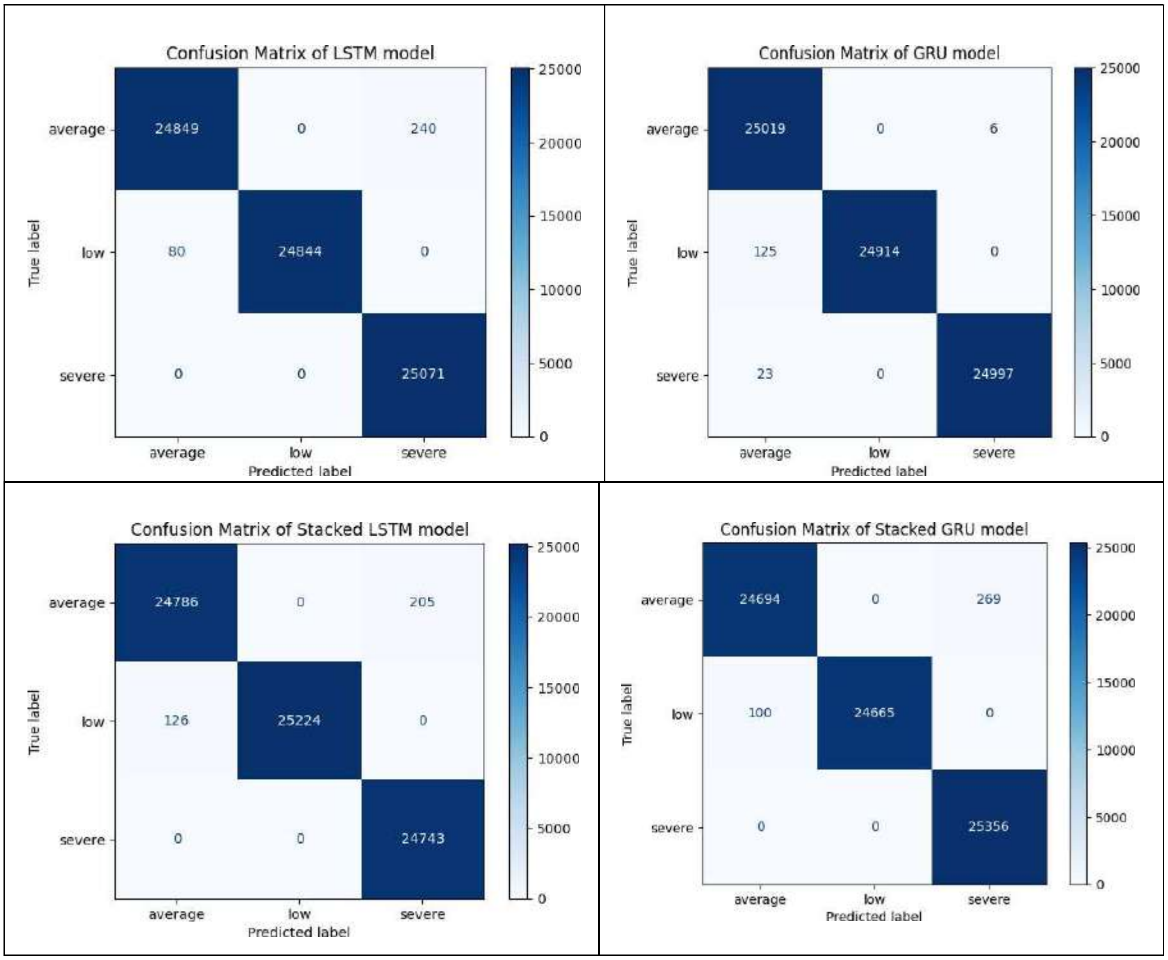


Figure 7: Confusion matrix depicting the performance of models on Data-I for three air quality classes: Low, Average, and Severe

Table 5 presents the performance metrics of models across three different classes of smart home dataset i.e. Low, Average, and Severe based on precision, recall, and F1 -score, to evaluate their effectiveness in classifying them.

Table 5: Class-wise analysis of models for Data-I to show their precision, recall, and F1- scores across Low, Average, and Severe air quality classes

Models	**Class**	**Precision**	**Recall**	**F1score**
MLP	**Low**	1.00	0.9918	0.9958
	Average	0.9751	1.00	0.9873
	Severe	1.00	0.9829	0.9913
RNN	**Low**	1.00	0.9980	0.9989
	Average	0.9980	0.9946	0.9962
	Severe	0.9945	1.00	0.9972

ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

Bidirectional GRU	**Low**	1.00	0.9951	0.9975
	Average	0.9951	0.9995	0.9972
	Severe	0.9991	1.00	0.999
Bidirectional LSTM	**Low**	1.0	0.9965	0.9982
	Average	0.9965	0.9951	0.9958
	Severe	0.9951	1.0	0.9975
LSTM	**Low**	1.00	0.9967	0.9983
	Average	0.9967	0.9904	0.9935
	Severe	0.9905	1.00	0.9952
GRU	**Low**	1.00	0.9948	0.9973
	Average	0.9941	0.9997	0.9968
	Severe	0.9997	0.9997	0.9997
Stacked LSTM	**Low**	1.00	0.9950	0.9974
	Average	0.9949	0.9917	0.9932
	Severe	0.9917	1.00	0.9958
Stacked GRU	**Low**	1.00	0.9959	0.9979
	Average	0.9959	0.9892	0.9925
	Severe	0.9895	1.00	0.9947

After analyzing the table, it has been observed that in case of ***Low class***, all the models have performed well by scoring the perfect precision scores of 1.00 which means that every positive prediction made by the model is indeed a true positive. But with slight variations among recall and F1 score. This variation indicates models are identifying true positives with differing success rates. Some models might be missing more true positives (higher false negatives) than others. For this class, the highest recall and F1 score value has been computed by RNN with 0.9980 and 0.9989 followed by LSTM and bidirectionalLSTM with 0.9967 and 0.9965 as recall and 0.9983 and 0.9982 as F1 score each. However, MLP shows the high number of false negatives and misses many true positives by computing the least recall and F1 score values with 0.9918 and 0.9958 respectively. In case of ***Average class*** of air quality level, only MLP computed the perfect recall of 1.00 as compared to other models which indicates that MLP correctly identifies all the true positive cases without missing any. While examining the models for rest of the parameters such as precision, the highest value has been computed by RNN with 0.9980 followed by LSTM and BidirectionalLSTM with 0.9967 and 0.9965 respectively which defines the true positive prediction made by them. Likewise, for recall and F1 score, the highest values have been computed by GRU with 0.9997 and 0.9968 as well as BidirectionalGRU with 0.9995 and 0.9972 respectively. But on the contrary, the lowest values for all the performance metrics has been obtained by GRU in case of precision with 0.9941, stacked GRU in terms of recall of F1 score with 0.9892 and 0.9925 respectively. At the end, for ***Severe class,*** only MLP computed the perfect precision score while as the rest of the models such as RNN, BidirectionalGRU, BidirectionalLSTM, LSTM, StackedLSTM, and StackedGRU obtained the perfect recall scores of 100%. GRU maintained the balance relationship between the metrics

ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

by computing 0.9997 each which means very low number of false positives and false negatives of severe class. In addition to this, Stacked GRU obtained the lowest precision value of 0.9895 and MLP had least recall and F1 score value which defines there is still a room for improvement.

Analysis of models for the Data II of Room 776

Table 6 presents a comparative analysis of various deep learning models based on their training accuracy and loss for the data collected from room 776.

Table 6: Accuracy and loss values for training and validation phases of neural network models applied to Data-II

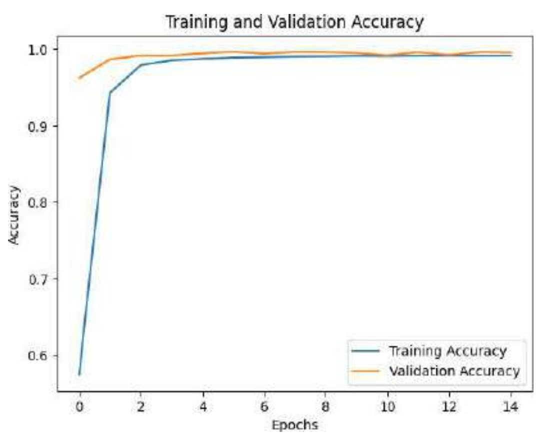
Models	Training		Validation	
	Acc	Loss	Acc	Loss
MLP	99.17	0.0276	99.58	0.0140
RNN	99.26	0.0240	99.43	0.0158
Bidirectional GRU	99.27	0.0238	99.41	0.0151
Bidirectional LSTM	99.14	0.0273	99.45	0.0130
LSTM	99.20	0.0255	99.47	0.0143
GRU	99.24	0.0247	99.63	0.0118
Stacked LSTM	99.13	0.0281	99.64	0.0117
Stacked GRU	99.16	0.0264	99.34	0.0177

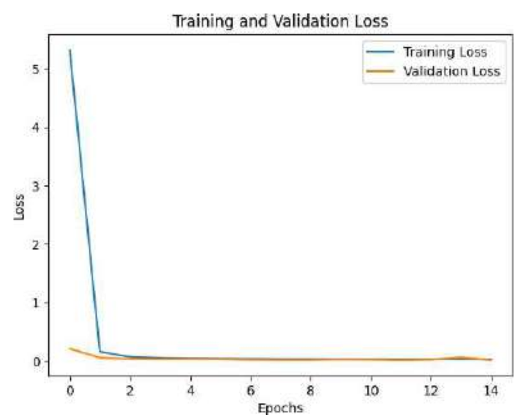
In the case of ***training phase***, the Bidirectional GRU achieves the highest training accuracy of 99.27%, closely followed by the RNN with 99.26%, and the GRU with 99.24%. This suggests that these recurrent-based models are particularly effective in capturing the patterns in the training data. In terms of training loss, which indicates how well the model fits thetraining data (with lower values being better), the Bidirectional GRU also performs the best, with the lowest loss of 0.0238. This is followed by the RNN with a loss of 0.0240, and the GRU with 0.0247. These low loss values, in conjunction with their high accuracies, suggest that these models not only learn well but also generalize effectively on the training data without overfitting. On the other hand, models like the Stacked LSTM and Stacked GRU, while still achieving high accuracies (99.13% and 99.16% respectively), show slightly higher loss values (0.0281 and 0.0264 respectively) compared to their simpler counterparts. This could imply that the added complexity of stacking layers does not necessarily translate to better performance on the training data and may even lead to marginally increased training loss. From the ***validation data***, all models show high performance with accuracies ranging from 99.34% to 99.64%, and low loss values between 0.0117 and 0.0177. The Stacked LSTM model achieves the highest validation accuracy at 99.64%, closely followed by the GRU with 99.63%, indicating that these models generalize particularly well to unseen data. This high performance suggests that these models effectively capture the underlying patterns in the dataset. In terms of validation loss, which measures how well the model fits the validation data (with

ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

lower values being better), the Stacked LSTM again performs the best with the lowest loss of 0.0117, closely followed by the GRU with a loss of 0.0118. These low loss values reinforce the superior generalization capability of these models. Interestingly, the MLP model, despite its simpler architecture, achieves a very high validation accuracy of 99.58% and a low loss of 0.0140, indicating that it is also a strong performer and can be considered a competitive model for this task. The RNN, Bidirectional GRU, and Bidirectional LSTM models show slightly lower validation accuracies (99.43%, 99.41%, and 99.45%, respectively) and higher losses (0.0158, 0.0151, and 0.0130, respectively) compared to the GRU and Stacked LSTM. This suggests that while they are still effective, they may not capture the validation data patterns as well as the GRU-based models and the Stacked LSTM. The Stacked GRU model obtained 99.34% as validation accuracy on a higher loss of 0.0177 which indicates the possibility of overfitting or difficulty in effectively training the deeper architectures.

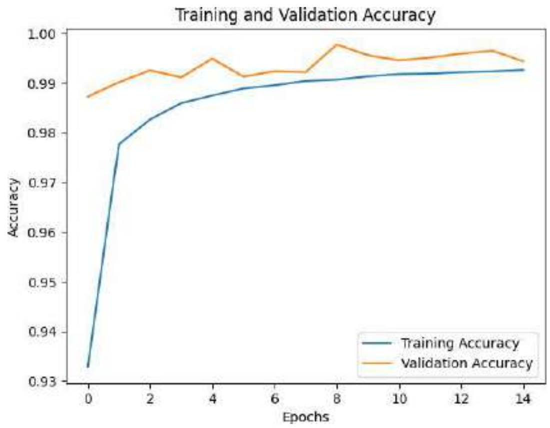
Like in the case of Data-I, here also the learning curves of the models for training and validation accuracy as well as loss for 15 epochs is defined in Figure 8. According to our research, it has been observed that MLP and RNN models exhibit a good fit of learning curves while as in case of Bidirectional GRU, a peak can be seen which defines that the model might not have learned enough to make accurate predictions but as the training progresses, the model learns to capture more complex patterns in the data which leads to its improved performance.

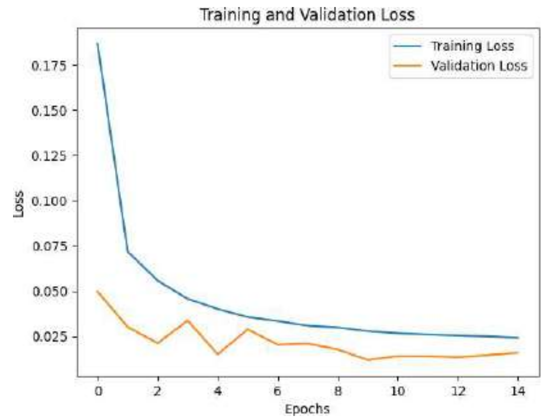




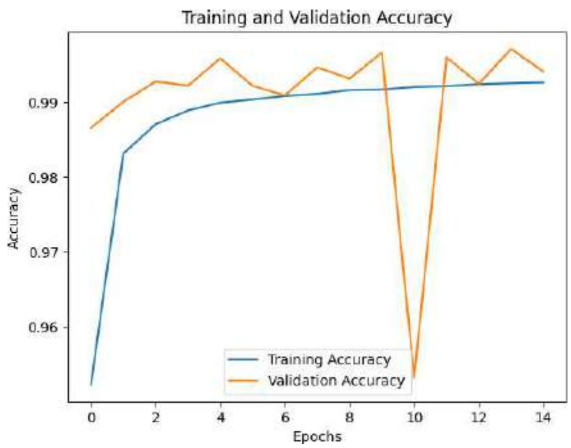
ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

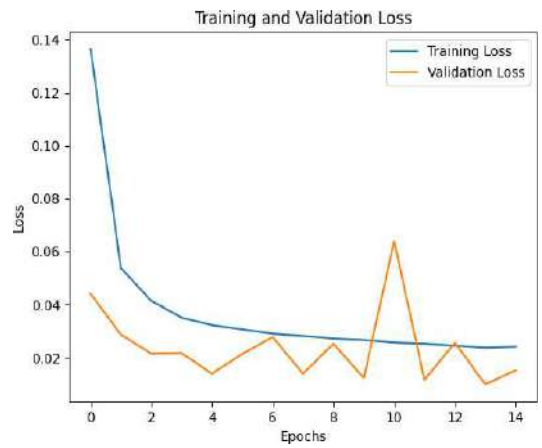
MLP



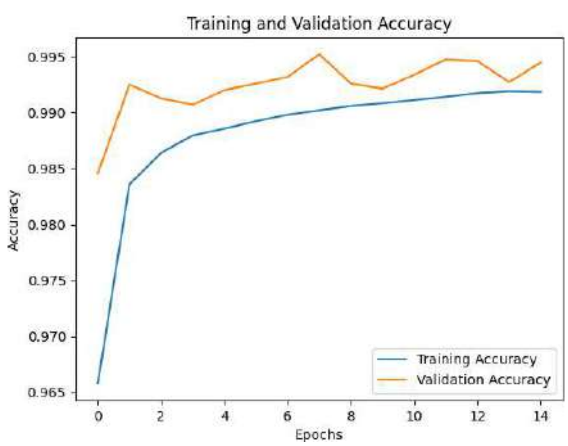


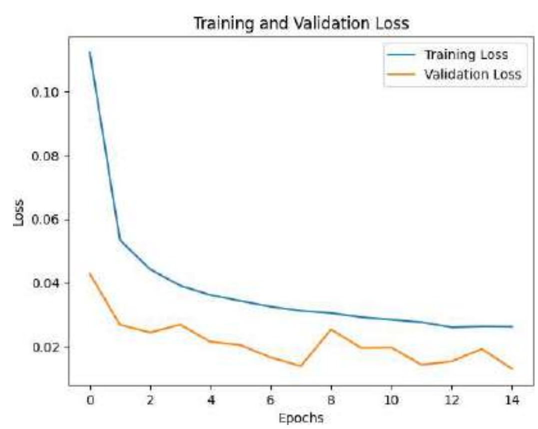
RNN





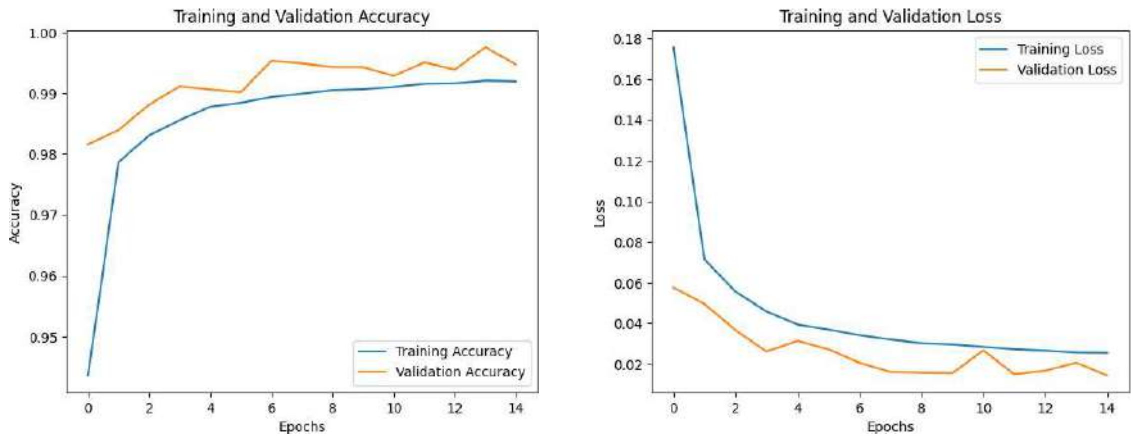
Bidirectional GRU



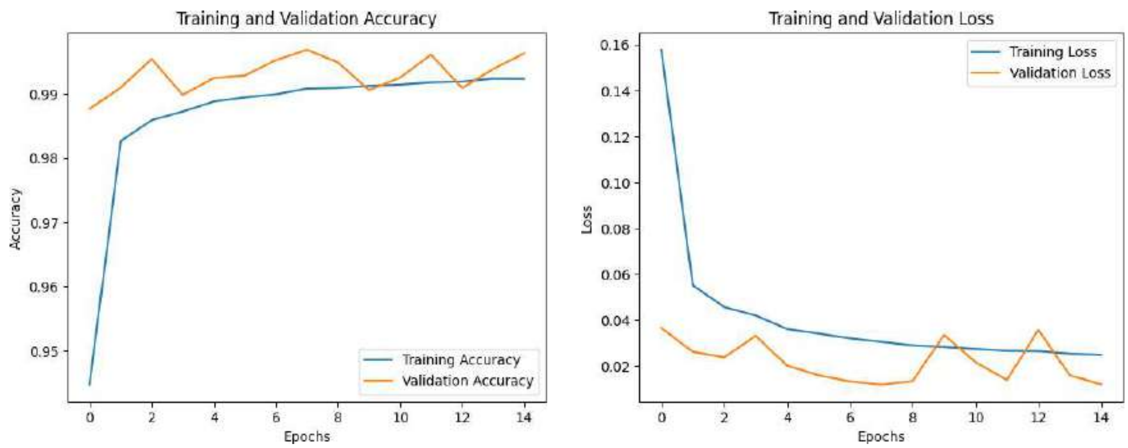


Bidirectional LSTM

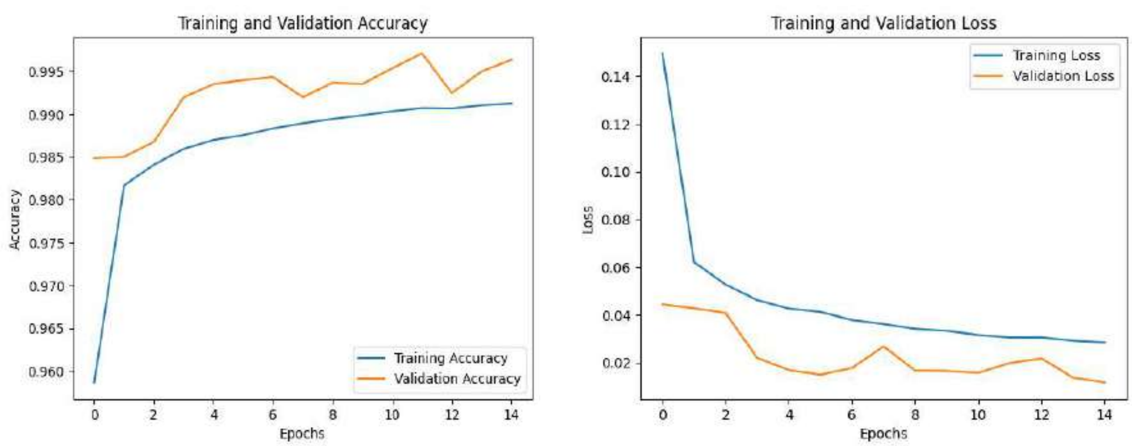
ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)



LSTM

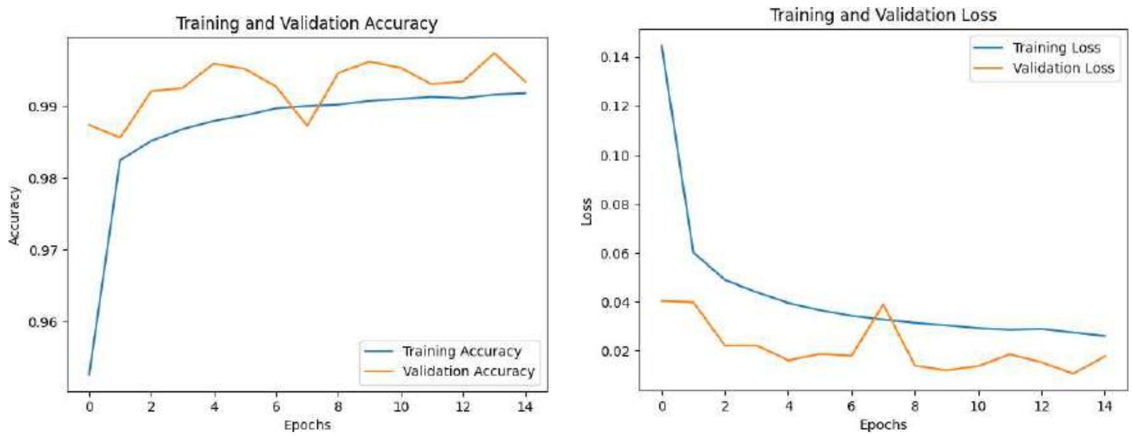


GRU



Stacked LSTM

ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)



Stacked GRU

Figure 8: Learning curves for models trained on Data-II to highlight differences in performance between training and validation dataset

Table 7 analyzes the performance of various deep learning models based on their precision, recall, and F1 score.

Table 7: Analysis of models for the energy efficient home appliances Data-II

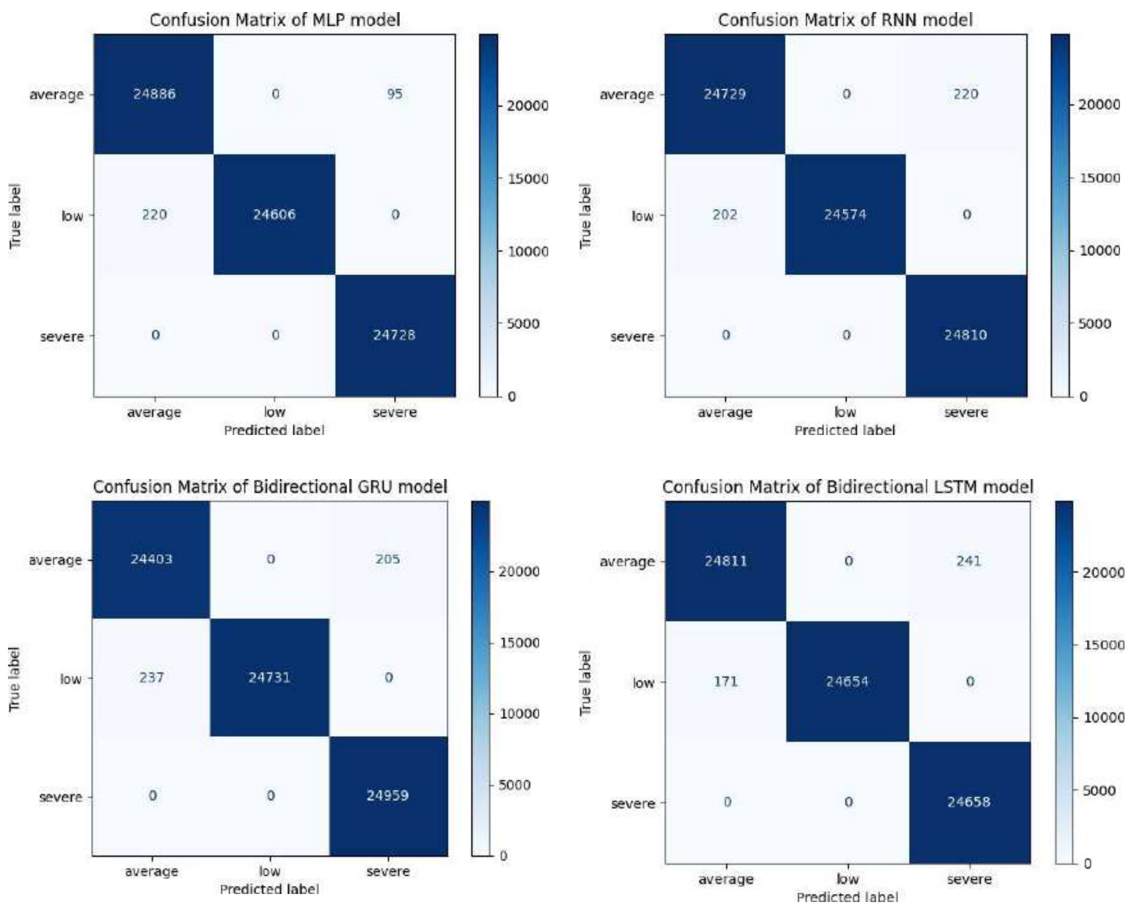
Models	**Precision**	**Recall**	**F1score**
MLP	0.9957	0.9957	0.9957
RNN	0.9943	0.9943	0.9942
Bidirectional GRU	0.9940	0.9940	0.9939
Bidirectional LSTM	0.9945	0.9945	0.9945
LSTM	0.9920	0.9973	0.9946
GRU	0.9963	0.9963	0.9962
Stacked LSTM	0.9959	0.9964	0.9961
Stacked GRU	0.9961	0.9966	0.9949

GRU and Stacked GRU indicated the effectiveness in minimizing the false positives by obtaining the highest precision values of 0.9963, 0.9961, and 0.9959 respectively followed by StackedLSTM and MLP with their scores as 0.9959 and 0.9957. On the other hand, in terms of recall, LSTM with highest value of 0.9973 reflected its ability in identifying true positives albeit of having lower precision of 0.9920 which ultimately computed 0.9946 as F1 score. Apart from LSTM, Stacked GRU, Stacked LSTM, and GRU also presented their efficiency in predicting the actual instances by obtaining the top recall values with 0.9966, 0.9964, and 0.9963 on an F1 score of 0.9949, 0.9961, and 0.9962 respectively. MLP model also demonstrates its

ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

robustness by having a good F1 score of 0.9957 while as Bidirectional GRU and BidirectionalLSTM generated the lowest values of precision and recall with 0.9940 and0.9945 each. It suggested that the models may not perform as consistently well as the other models.

Here also, a confusion matrix of 3x3 size has been also created to analyze the actual and predicted values of the models as well as to get a concise view of how models are performing across the class, shown in Figure 9.



ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

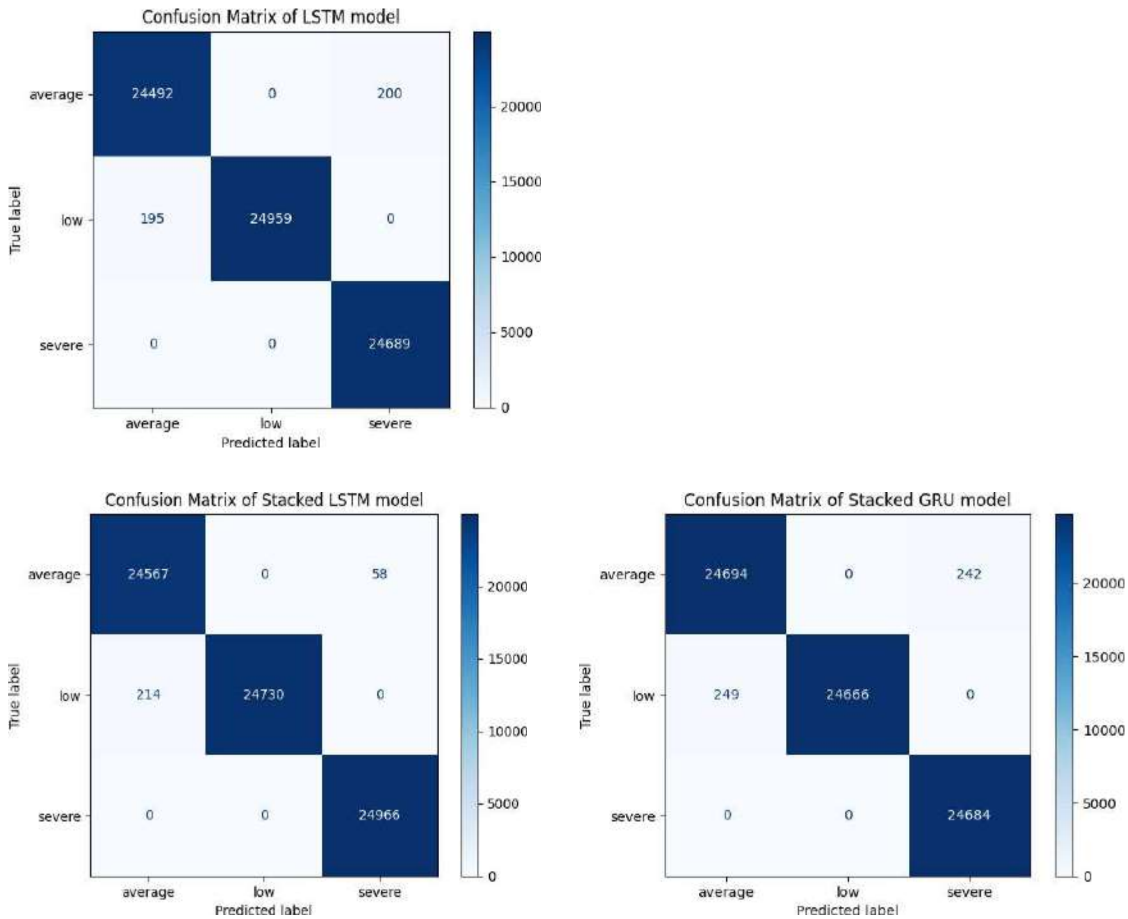


Figure 9: Confusion matrix showing model predictions versus actual values for Data-II across air quality levels: Low, Average, and Severe

Table 8 analyzes the performance of applied classifiers based on three different classes of dataset i.e. Low, Average, and Severe for the evaluation metrics as precision, recall, and F1 score.

Table 8: Evaluation of model performance in distinguishing air quality levels (Low, Average, Severe) in Data-II

Models	Class	Precision	Recall	F1score
MLP	**Low**	1.00	0.9911	0.9955
	Average	0.9912	0.9961	0.9936
	Severe	0.9961	1.00	0.9980
RNN	**Low**	1.00	0.9918	0.9958
	Average	0.9918	0.9911	0.9914
	Severe	0.9912	1.00	0.9955
Bidirectional GRU	**Low**	1.00	0.9905	0.9952
	Average	0.9903	0.9916	0.9909
	Severe	0.9918	1.00	0.9958

Bidirectional LSTM	**Low**	1.00	0.9931	0.9966
	Average	0.9932	0.9904	0.9918
	Severe	0.9903	1.00	0.9951
LSTM	**Low**	0.9922	1.00	0.9961
	Average	0.9921	0.9919	0.9919
	Severe	0.9919	1.00	0.9959
GRU	**Low**	1.00	0.9913	0.9956
	Average	0.9914	0.9977	0.9945
	Severe	0.9976	1.00	0.9987
Stacked LSTM	**Low**	1.00	0.9914	0.9957
	Average	0.9913	0.9979	0.9945
	Severe	0.9964	1.00	0.9982
Stacked GRU	**Low**	1.00	0.9900	0.9949
	Average	0.9898	0.9999	0.9948
	Severe	0.9902	1.00	0.9950

For the ***low*** class, all models indicated their robustness in perfectly classifying as well as identifying the classes of them obtained perfect precision score of 100% except LSTM as this model computed the perfect recall score. However, there are also few classifiers which computed lowest recall values like MLP (0.9911), RNN (0.9918), BidirectionalGRU (0.9905), GRU (0.9913), StackedLSTM (0.9914), and StackedGRU (0.9900) which indicates that these models are failing to correctly identify a significant number of true positive instances for this particular class. In addition to this, Bidirectional LSTM show slight variations with a recall of 0.9931 which results in F1 scores of 0.9966, depicts slight better performance compared to others. Likewise, in case of ***average*** class, all the models except StackedGRU maintain high precision, recall, and F1 scores, typically above 0.99. However, models like the Bidirectional GRU, RNN, LSTM, and Bidirectional LSTM have slightly lower F1 scores of 0.9909, 0.9914, 0.9919, and 0.9918, respectively, due to their slight fall of performance in either precision or recall. Stacked GRU shows highly sensitive nature in detecting the instances of average class by computing the highest recall value of 0.9999 but on a lower precision score of 0.9898 which ultimately leads to an F1 score of 0.9948. Similarly, lowest recall values have been also computed by the models such as RNN (0.9911), BidirectionalGRU (0.9916), BidirectionalLSTM (0.9904), and LSTM (0.9919) which means that they are not able to classify the actual positive values correctly. In the ***severe*** class, all the models again show very high performance by achieving perfect recall score of 1.00. This characteristic indicates that the models are able to identify all instances of the severe class correctly. In case of F1 score, GRU and Stacked LSTM models stand out by computing the highest values as 0.9987 and 0.9982 respectively, while as the other models like MLP and RNN obtained slightly lower F1 scores (0.9980 and 0.9955) each. In addition to this, Stacked GRU and Bidirectional LSTM obtained the lowest precision score of 0.9902 and 0.9903 which indicates these models are not able to predict the true positive

ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

classes and require room for improvement.

Table 9 provides the detailed analysis of the time frame taken by different deep learning models to process Data I and Data II. The Multilayer Perceptron because of its simple architecture stands out with the shortest training time of 1 hour which indicates its faster computational efficiency compared to the other models. On the other hand, due to their deeper and complex architectures, the Stacked LSTM and Stacked GRU require longer training times of 1 hour and 40 minutes and 2 hours, respectively. Similarly, the Bidirectional LSTM and Bidirectional GRU models computed in the longest training times of 2 hours and 5 minutes and 2 hours, respectively. While as the RNN model, LSTM and GRU models trained in an average period of time with a training time of 1 hour and 30 minutes, 1 hour and 25 minutes and 1 hour and 40 minutes respectively.

Table 9: Overall execution time of applied learning models for Data-I and Data-II, to underscore the impact of model complexity on computational efficiency

Models	Time frame
LSTM	1 hour 25 min
GRU	1 hour 40 min
Stacked LSTM	1 hour 40 min
Stacked GRU	2 hour
MLP	1 hour
RNN	1 hour 30 min
Bidirectional GRU	2 hour
Bidirectional LSTM	2 hour 5 min

Overall, in terms of practical application scenarios, the excellent performance of Bidirectional GRU and Stacked LSTM models in terms of accuracy and loss could be applied to real-time energy optimization systems in smart homes. These models could dynamically adjust energy usage for heating, ventilation, and lighting based on indoor air quality parameters such as CO2 levels and temperature, ensuring energy efficiency while maintaining comfort. Additionally, insights from the learning curves, which highlight fluctuations in training and validation accuracy, can guide practical decisions in hyperparameter tuning during the deployment of models in real-world scenarios. The SMOTE-ENN technique for class balancing could be effectively applied in homes with uneven energy consumption patterns, such as those with varying seasonal appliance usage, ensuring that models handle such imbalances for accurate predictions. Lastly, the execution times of different models, detailed in Table 8, could inform decisions in scenarios where computational efficiency is crucial, such as in real-time energy management systems requiring rapid model updates and predictions. These practical applications demonstrate how deep learning models can be directly translated into impactful, energy-saving solutions for residential buildings, contributing to sustainability and efficiency.

ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

III. CONCLUSION

This paper highlighted significant advancements in applying deep learning techniques to improve energy efficiency in home appliances. By analyzing factors such as CO2 levels, humidity, and temperature, the models demonstrated impressive results in optimizing energy consumption and promoting sustainability. Among the classifiers tested, Bidirectional GRU and Stacked LSTM outperformed others in terms of accuracy and loss for data collected from two rooms 415 and 776, showcasing the potential of AI-driven approaches to revolutionize energy management in smart homes. Furthermore, these models can optimize real-time decision-making in smart appliances, contributing to significant energy cost savings and a reduced environmental footprint. This research highlights the broader potential for integrating AI-driven approaches into energy policies and sustainability strategies, enabling more effective reductions in residential energy consumption and combating climate change However, the research's dependency on specific datasets and observed fluctuations during training and validation point to limitations like overfitting and constrained generalizability. Addressing these issues requires diversifying datasets, adjusting learning rates, increasing batch sizes, or employing regularization techniques to stabilize training and enhance model robustness. Future research should also explore advanced optimization methods like Adaptive Moment Estimation, Root Mean Square Propagation, and evolutionary algorithms. Additionally, integrating IoT devices and real-time data processing can enhance responsiveness and scalability, bridging the gap between technological advancements and their practical deployment in sustainable energy systems.

Authors' Contribution:

Writing — original draft, J.S.S., and S.A.; Methodology, J.S.S., S.A. and S.K.; Formal analysis, J.S.S., and S.A.; Analysis result review, J.S.S., S.A. and S.K. All authors have read and agreed to the published version of the manuscript.REFERENCES

[1] Santamouris, M., and K. Vasilakopoulou. ―Present and future energy consumption of buildings: Challenges and opportunities towards decarbonisation‖. e-Prime-Advances in Electrical Engineering, Electronics and Energy, 1 (2021): 100002.

[2] Olatunde, Tosin Michael, Azubuike Chukwudi Okwandu, and Dorcas Oluwajuwonlo Akande ―Reviewing the impact of energy-efficient appliances on household consumption‖, International Journal of Science and Technology Research Archive, 2024, 06(02), 001–011 (2024).

[3] ―Energy Efficiency: Buildings and Industry‖. Energy.gov,www.energy.gov/eere/energy- efficiency-buildings-and-industry.

[4] Chen, Wei, Alharthi, Majed, Zhang, Jinjun, Khan, Irfan, ―The need for energy efficiency and economic prosperity in a sustainable environment‖, Gondwana Research, 127 (2024): 22-35.

[5] Karatzas, S., Merino, J., Puchkova, A., Mountzouris C., Protopsaltis G., Gialelis, J., Parlikad, A.K., ―A virtual sensing approach to enhancing personalized strategies for indoor environmental quality and residential energy management‖, Building and Environment, Volume 261, 2024: 111684.

[6] Verma, Aditya, and Yogesh Kumar. ―Study on machine learning based energy efficiency in developed countries‖. *2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC)*. IEEE, 2020.

[7] Abobakirov, A. ―Energy Efficient Building Materials in the Design of Buildings and Structures‖. *HOLDERS OF REASON* 1.3 (2023): 406-412.

[8] Khalil, M., McGough, A.S., Pourmirza, Z., Pazhoohesh, M., Walker, S., ―Machine Learning, Deep Learning and Statistical Analysis for forecasting building energyconsumption – A systematic review, Engineering Applications of Artificial Intelligence, Volume 115, 2022.

[9] Yong, X., Yonghua, C., Jiaojiao, X., and Zheyou, C., ―Research on sustainability evaluation of green building engineering based on artificial intelligence and energy consumption‖, *Energy Reports,* 8, 2022 : 11378-11391.

[10] Park, S., ―Machine Learning-Based Cost-Effective Smart Home Data Analysis and Forecasting for Energy Saving‖. *Buildings* 13.9 (2023): 2397.

[11] Morteza, A., Yahyaeian, A.A., Mirzaeibonehkhater, M., Sadeghi, S., Mohaimeni, A., Taheri, S., ―Deep learning hyperparameter optimization: Application to electricity and heat demand prediction for buildings‖, Energy and Buildings, 289, 2023: 113036.

[12] Khan, Junaid, et al. ―A Hybrid Machine Learning and Optimization Algorithm for Enhanced User Comfort and Energy Efficiency in Smart Homes‖. (2024).

[13] Shree, Lakshmi, et al. ―Efficient Optimization of Energy Consumption at Home through Machine Learning‖. *2024 5th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI)*. IEEE, 2024.

[14] Khan, Murad, Junho Seo, and Dongkyun Kim. ―Towards energy efficient home automation: a deep learning approach‖. *Sensors* 20.24 (2020): 7187.

ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

[15] https://www.kaggle.com/datasets/ranakrc/smart-building-system

[16] Kasaraneni, Purna Prakash, et al. ―Machine learning-based ensemble classifiers for anomaly handling in smart home energy consumption data‖. *Sensors* 22.23 (2022): 9323.

[17] Yoon, Young Ran, et al. ―A non-intrusive data-driven model for detailed occupants' activities classification in residential buildings using environmental and energy usage data‖. *Energy and Buildings* 256 (2022): 111699.Ding, Yong, Lingxiao Fan, and Xue Liu. ―Analysis of feature matrix in machine learning algorithms to predict energy consumption of public buildings‖. *Energy and Buildings* 249 (2021): 111208.

[18] Alharbi, Amal H., et al. ―Forecasting of energy efficiency in buildings using multilayer perceptron regressor with waterwheel plant algorithm hyperparameter‖. *Frontiers in Energy Research* 12 (2024): 1393794.

[19] Fan, Cheng, et al. ―Assessment of deep recurrent neural network-based strategies for short-term building energy predictions‖. *Applied energy* 236 (2019): 700-710.

[20] Natarajan, Yuvaraj, et al. ―Enhancing Building Energy Efficiency with IoT-Driven Hybrid Deep Learning Models for Accurate Energy Consumption Prediction‖. *Sustainability* 16.5 (2024): 1925.

[21] Singh, Sushil Kumar, et al. ―GRU-based digital twin framework for data allocation and storage in IoT-enabled smart home networks‖. *Future Generation Computer Systems* 153 (2024): 391-402.

[22] Attarde, Khush, and Javed Sayyad. ―A CNN and BiLSTM Fusion Approach Toward Precise Appliance Energy Forecasts‖. *International Journal of Intelligent Engineering & Systems* 17.3 (2024).

[23] Niu, Dongxiao, et al. ―Short-term multi-energy load forecasting for integrated energy systems based on CNN-BiGRU optimized by attention mechanism. *Applied Energy* 313 (2022): 118801‖.

[24] Alghamdi, Mona Ahamd, S. Abdullah, and Mahmoud Ragab. ―Predicting Energy Consumption Using Stacked LSTM Snapshot Ensemble‖. *Big Data Mining and Analytics* 7.2 (2024): 247-270.

[25] Han, Tao, et al. ―An efficient deep learning framework for intelligent energy

ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

management in IoT networks‖. *IEEE Internet of Things Journal* 8.5 (2020): 3170-3179.Izonin, Ivan, et al. ―Machine learning for predicting energy efficiency of buildings: a small data approach‖. *Procedia Computer Science* 231 (2024): 72-77.

[26] Long, Luong Duc. ―An AI-driven model for predicting and optimizing energy-efficient building envelopes‖. *Alexandria Engineering Journal* 79 (2023): 480-501.

[27] Mayer, Kevin, et al. ―Estimating building energy efficiency from street view imagery, aerial imagery, and land surface temperature data‖. *Applied Energy* 333 (2023): 120542.

[28] Kapoor, Nitika, and Yogesh Kumar. ―The Efficient Management of Renewable Energy Resources for Vanet-Cloud Communication‖. *Nature-Inspired Computing Applications in Advanced Communication Networks*. IGI Global, 2020. 228-253.

[29] https://www.greenbuilt.org/energysaversnetwork/home-energy-saving-tips/

[30] Liao, Qi, et al. ―Deep learning for air quality forecasts: a review‖. *Current Pollution Reports* 6 (2020): 399-409.