Svedberg Open
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

An Intelligent Machine Learning Framework For Autonomous Control of Conventional Air Conditioners

Shubhangi Mathe[1] , Pooja Kulkarni[2]

[1]Department of Computer Science, Vishwakarma University, Kondhwa, Pune, India. Email: shubhangi.mathe-438@vupune.ac.in
[2]Department of Computer Science, Vishwakarma University, Kondhwa, Pune, India

ABSTRACT

In recent years, the demand for smart air conditioners (ACs) has expanded swiftly due to the need for comfort, automation, and energy efficiency. However, upgrading conventional AC units to smart systems is often expensive. To overcome this, the present research proposes a cost-effective IoT-based device that transforms a non-smart AC into an intelligent, smart AC. The developed model automatically adjusts cooling based on room temperature, occupancy, humidity, and CO_2 concentration without requiring a remote control or mobile application. The idea behind this research is to design and implement a low-cost, power-saving system that functions automatically without human intervention. The model is built using an ESP32 microcontroller interfaced with two proximity sensors, two temperature sensors, two humidity sensors, and one CO_2 sensor. One temperature and humidity sensor pair are placed outside the room, while the CO_2 sensor is located inside. The system monitors these parameters in real time and utilizes a logistic regression algorithm to predict whether high cooling is necessary. When occupancy exceeds four persons and environmental factors indicate increased CO_2 and humidity, the device automatically reduces the AC's temperature to an optimal value. Experimental results give up to 82.72% accuracy in occupancy detection and improved cooling efficiency under crowded conditions. This economically feasible model provides a practical solution for domestic environments, senior citizens, and educational institutions, enabling intelligent control of traditional AC units without manual operation. This model is useful in conference halls of school or colleges. Also useful for senior citizens on domestic platforms.

Conventional air conditioners are widely used in residential and institutional environments, but most existing units require continuous manual operation and do not have the sensing and decision-making capabilities of modern smart air-conditioning systems. This study introduces an intelligent machine learning framework that enables conventional air conditioners to perform autonomous cooling control using an embedded sensing and computing platform. The proposed framework employs an ESP32 microcontroller to acquire real-time information from temperature, humidity, CO_2, and occupancy sensors. Environmental parameters are monitored from the indoor and outdoor surroundings to identify changes in thermal conditions and occupancy-related cooling requirements. The collected data are processed at the edge, allowing the system to make control decisions without depending on continuous user interaction or a mobile application. Logistic Regression is incorporated as the machine learning model to classify the required cooling condition based on the monitored parameters. When occupancy and environmental conditions indicate a higher cooling requirement, the system automatically modifies the operating temperature of the conventional AC. Experimental testing of the developed prototype achieved an occupancy detection accuracy of 82.72% and demonstrated reliable automatic adjustment of cooling under different occupancy conditions. The proposed framework provides a low-cost approach for extending intelligent control capabilities to existing conventional air conditioners without replacing the complete AC unit. The system can be applied in classrooms, conference halls, domestic environments, and other shared spaces where autonomous temperature regulation and reduced dependence on manual AC operation are desirable.

Keywords: Machine Learning, Intelligent Control, Conventional Air Conditioner, Autonomous Operation, ESP32, Logistic Regression, Occupancy Detection, Temperature Regulation. IoT, ESP32, smart air conditioning, logistic regression, occupancy detection, energy efficiency.

INTRODUCTION

In the past few years, the advancement of the Internet of Things (IoT) has resulted in the creation of numerous smart home applications designed to enhance comfort, automation, and energy efficiency.

Among these, smart air conditioners (ACs) have acquired significant focus because of their capacity to optimize energy consumption based on user behaviour and environmental conditions. Recently, many of the homes are using smart air conditioners. Smart air conditioners are used for enhanced comfort, convenience, and energy efficiency through remote control via a smartphone app or voice command, scheduling, and geofencing. They can adapt to user convenience and room conditions. However, replacing conventional AC units with smart ones involves high costs, making it impractical for most households and institutions. AC room temperature can be controlled with the help of mobile application. It can be done from anywhere as these smart ACs are Wi-Fi enabled. Conventional ACs do not have such features. It can be operated with the help of remote control only. Conventional ACs consume more electricity due to their fixed-speed operation and lack of automation. Replacing conventional AC with smart AC would increase the expenses 100% more, along with the installation overhead. To overcome this limitation, this research proposes an IoT based model that fully converts non-smart AC into smart one. The developed model monitors various environmental parameters such as temperature, humidity, carbon dioxide (CO_2) concentration, and room occupancy to control the AC operation automatically. Unlike existing solutions that rely on smartphone applications or remote controls, the proposed system functions independently and can be deployed in homes, schools, and institutions where automation can significantly improve comfort and energy efficiency.

The major contribution of this work lies in the integration of multiple sensors with an ESP32 microcontroller and the application of a logistic regression model for intelligent decision-making. This approach ensures energy optimization by adjusting the AC temperature dynamically based on real-time environmental and occupancy data. Various studies are done on home automation. Khalil Mohamed et. al. in 2023 published their work in the International Review of Automatic Control. The author proposed the alternative design of a home automation system. A prototype was presented that integrates IoT with cloud computing. Node MCU serves as its central control unit. This system functions in both manual and remote modes and combines a number of smart gadgets with security alerts [1]. Alma Rosa Mena et. al. in 2022 in the journal of MDPI said that sensors play a very important role in IoT. Use of appropriate sensors makes IoT devices more flexible and reliable. The author reviewed various studies on the sensors[2]. Abdul Lateef Haroon et. al. in 2024 presented the work in IEEE journal, stating that home automation supports modern living by providing ease of user centric design and comfort. IoT technology can be fused with sensors, microcontrollers and smartphone applications. Remote monitoring can be fulfilled by using Wi-Fi connectivity[3] . Ullo et. al. in the year 2021, studied and published the work in MDPI journal the monitoring of air quality, water quality, rapid pollution and agriculture systems. With the growth of contemporary sensors and the internet of things, smart environment monitoring has become possible. Various SEM (smart environment monitoring) studies focus on various machine learning algorithms, advancements in sensor as well as wireless technology[4]. Raspberry Pi, ESP32, Arduino are the trending microcontrollers used on a large scale in IoT technology. Microcontroller embedding with the temperature sensor can be used to note down the current temperature readings. Budijono et.al. in the year 2021 put the idea of ESP32 based intelligent temperature detection, and published the work in the IOP journal. It is used to keep an eye on the freezer's operation, measure the temperature continually for a predetermined amount of time, save the measurement data, and compute the results for preventive maintenance tasks. In many of the food outlets, freezer conditions need to be checked for a certain temperature threshold value. It can be achieved through the use of ESP32 with temperature sensors [5] . F A Tsvetanov in 2021 presented the work of design of a model that records the live data. This work is published in IOP publishing. The concepts of TSBD (Time series Database) and cloud storage database. These databases are useful for the performance analysis of the designed IoT model. ESP32 microcontroller embedded with various sensors, generating vast amounts of environmental data. This live data with each time stamp can be logged on the cloud databases[6]. Mark J. Mendell in the year 2024 in the Springer Nature, put the review of the impact of CO2 level on indoor air quality. With the support of IAQ (indoor air quality) guidelines, a total of 43 guidelines were identified. From this useful study, the most common CO2 limit identified as 1000 ppm[7]. Marik Babiuch and others in 2019 put the practical approach in the Institute of Electrical and Electronics Engineers. This approach is towards testing various ESP32 versions for the establishment of the WIFI module and to record the sensor readings. Wi-fi connected status is displayed with the help of OLED and for better results ESP32 Wrover development board is utilized [8]. Roberto Pashic and others have done various studies and published the work in the TEMEL-ij international journal. The author focuses on successful wireless data logging. Many of the research papers suggest the ESP32 for data acquisition for different sensors. This study proposed that the ESP32 is a dual-core MCU with built-in WiFi and Bluetooth, 240MHz clock frequency, 512KB RAM, 36 pins, and a large range of available peripherals, including ADCs, DACs, UART, SPI, I2C, CAN2.0, PWM, and built-in temperature and hall effect sensors. Additionally, DOIT ESP32 DEVKIT V1 is among the greatest ESP32 MCU-based development boards[9]. Machine learning enhances operations of home automation by learning the user patterns and facilitates personalized and predictive control. Many of the researches proved that use of machine learning algorithms clarifies the performance of home automated appliances. Xin Zhou in 2021, published the comparison of different machine learning algorithms for predicting air-conditioning operating behaviour in open-plan offices. The way that air conditioners (AC) operate has a big impact on how much energy buildings use. Recent advances in machine learning (ML) techniques have demonstrated a high effectiveness in anticipating occupant. It is challenging to assess the advantages and disadvantages of various machine learning algorithms in this field of study due to a lack of comparison of these algorithms based on the same database and assessment indicators[10]. Using a state-of-the-art machine learning technique, this research aims to enhance home automation systems by including HAD based on sensors. R. Raghavendra in 2024 put the idea of Sensor-based human activity detection using a novel machine learning algorithm in home automation. Human activity detection using the Radiant Moth Flame-Optimized Probabilistic Decision Tree Algorithm (RMFO-PDTA) technique[11]. Malekzadeh (2023), EURASIP Journal, focused the study on performance of 5G networks using linear regression machine learning algorithm. Many of the studies proved that linear regression algorithms are very useful for the prediction of performance of home automation systems. It is also useful in other areas such as 5G networks[12]. Sathyanarayanan, S. in 2024, published a thorough understanding of the confusion matrix. Confusion matrix plays an important role for the evaluation of machine learning classification models. Along with the confusion matrix other factors such as F1 score, recall, precision, Cohen Kappa etc are also discussed [13]. Heydarian et. al. at 2022 in IEEE Access, detailed the study on Multiple Confusion matrix. Appropriate use of concise ML algorithms results in best performance. This study eliminates the ambiguity between false negative and false positive results[14].

1. METHODOLOGY

The proposed IoT-based AC control system was designed and implemented using the ESP32 microcontroller as the core processing unit. The framework of the model consists of the hardware setup reads the live data from the sensors. This data is stored on cloud storage and can be used to control the working of Air conditioner.

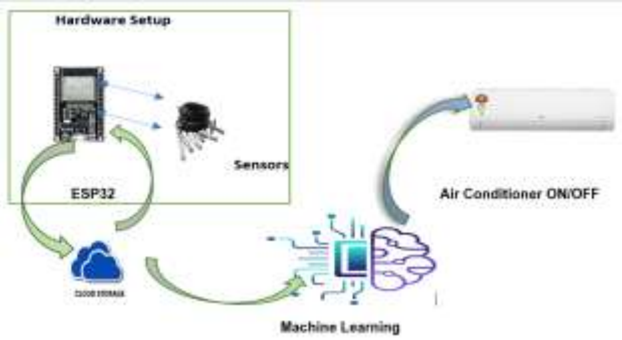

Fig. 1. Framework of the model

The hardware setup includes:

- **Two temperature sensors (DHT11)** – one placed inside and one outside the room to monitor thermal conditions.
- **Two humidity sensors** – one paired with each temperature sensor for comparative humidity analysis.
- **Two proximity sensors** – one positioned at the entrance to detect entry/exit events and another inside to estimate room occupancy.
- **One CO_2 sensor (MQ-135)** – to measure the concentration of carbon dioxide, indicating air quality and human presence.

The proposed IoT-based AC control system was designed and implemented using the ESP32 microcontroller as the core processing unit. The hardware setup includes:

- **Two temperature sensors (DHT11)** – one placed inside and one outside the room to monitor thermal conditions.
- **Two humidity sensors** – one paired with each temperature sensor for comparative humidity analysis.
- **Two proximity sensors (HC-SR04)** – one positioned at the entrance to detect entry/exit events and another inside to estimate room occupancy.
- **One CO_2 sensor (MQ-135)** – to measure the concentration of carbon dioxide, indicating air quality and human presence.

  The data from all sensors are collected in real time and transmitted to the ESP32 for processing. The microcontroller calculates the average temperature and humidity difference between indoor and outdoor environments. A logistic regression algorithm is then applied to predict whether a high cooling mode is necessary based on the following features:

- Room occupancy count
- Indoor temperature
- Indoor humidity
- CO_2 concentration

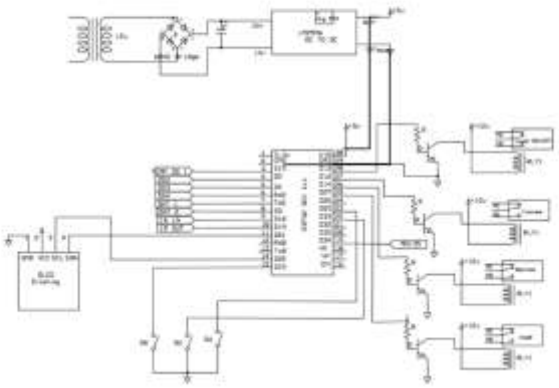

Fig. 2. Circuit diagram of the model

Fig 2 shows the circuit diagram of the proposed IoT-based smart AC control system. The ESP32 microcontroller forms the core of the hardware architecture and interfaces with all sensing and control modules. Temperature, humidity, and proximity sensors are connected to the ESP32 through digital and analog GPIO pins, enabling continuous monitoring of the room environment and occupancy status. An IR transmitter module, driven through a transistor, is used to send ON/OFF signals to the non-smart AC unit. A regulated 5V power supply powers the ESP32 and sensors. The circuit ensures seamless data acquisition, decision execution, and AC control as per the designed automation logic. Generally, if no one is in the room then 400 ppm of CO2 is available inside the room. One person exhibits 800 ppm. Likewise, if the occupancy exceeds four persons and the CO_2 or humidity level rises beyond the threshold say 3000 ppm, the model automatically signals the AC control relay to reduce the temperature. This threshold value can be set. At the experimental setup, given threshold values are set as: inside room temperature value greater than 25°C in combination with humidity greater than 80%. Conversely, when the room becomes unoccupied or environmental conditions does not cross the threshold value, the system switches the AC off to save energy. All collected data were logged on a cloud platform (Google Drive) for analysis and validation. Here, the role of machine learning comes into the picture. Random forest algorithm is useful to make the prediction and take the decisions to reduce or increase the AC temperature. The system's decision-making performance was evaluated by comparing predicted outcomes with actual user comfort levels.

Timestamp	AC Status	Outer Temp (°C)	Inner Temp (°C)	Outer Humidity (%)	Inner Humidity (%)	People Count
8/1/2025 13:03:41	ON	26	26.06	83.2	81.7	2
8/1/2025 13:04:43	ON	26	26.06	83.2	81.6	4
8/1/2025 13:05:43	OFF	25.97	26.03	83.2	81.6	0
8/1/2025 13:06:43	OFF	25.97	26.03	83.2	81.8	0
8/1/2025 13:07:42	OFF	25.97	26.03	83.3	82.2	0
8/1/2025 20:52:14	OFF	25.31	25.41	76.1	74.9	0
8/1/2025 20:53:13	OFF	25.25	25.28	80.1	78.1	0
8/1/2025 20:54:12	ON	25.25	25.34	80.3	78.3	1
8/1/2025 20:55:12	ON	25.25	25.28	80	78.3	1

Timestamp	AC Status	Outer Temp (°C)	Inner Temp (°C)	Outer Humidity (%)	Inner Humidity	People Count
8/2/2025 0:56:13	ON	25.25	25.28	83.5	81.7	4
8/2/2025 0:57:12	ON	25.25	25.28	83.2	81.6	4
8/2/2025 0:58:12	ON	25.22	25.25	83	81.3	4
8/2/2025 0:59:12	ON	25.22	25.28	82.9	81.3	4
8/2/2025 1:00:13	ON	25.22	25.25	82.7	81.1	4
8/2/2025 1:01:12	ON	25.25	25.25	82.5	81.1	4
8/2/2025 1:02:12	ON	25.25	25.25	82.4	81.1	4
8/2/2025 1:03:12	ON	25.25	25.25	82.4	80.9	4
8/2/2025 1:04:12	ON	25.22	25.25	82.2	80.7	4
8/2/2025 1:05:13	ON	25.19	25.19	82.3	80.5	4
8/2/2025 1:06:12	ON	25.09	25.16	80.4	78.7	4
8/2/2025 1:07:12	ON	24.94	25.16	78.8	77.5	4
8/2/2025 1:08:13	ON	24.84	25.06	77.6	76.6	4
8/2/2025 1:09:15	OFF	24.72	25	76.5	76	4
8/2/2025 1:10:13	OFF	24.59	24.97	75.8	75.2	4
8/2/2025 1:11:12	OFF	24.44	24.91	75	74.3	4
8/2/2025 1:12:12	OFF	24.41	24.81	74.4	73.8	4
8/2/2025 1:13:12	OFF	24.28	24.84	73.7	72.8	4

Fig. 3. Data set

The data set is created against one month. Inside room temperature and humidity, outside room temperature and humidity, reading of proximity sensors and CO2 levels were captured from the developed model.

2. RESULTS & DISCUSSIONS

After the data balancing, outlier detection can be done. It is a data point that behaves differently from all other data points in the dataset. Removal or treating of outliers gives us clean data for training purposes. If machine learning algorithms trained the data along with outliers then the model may distract and output different results.
Fig. 4 shows box plots for outliers detected in the data for indoor, outdoor temperature, indoor and outdoor humidity and people count.

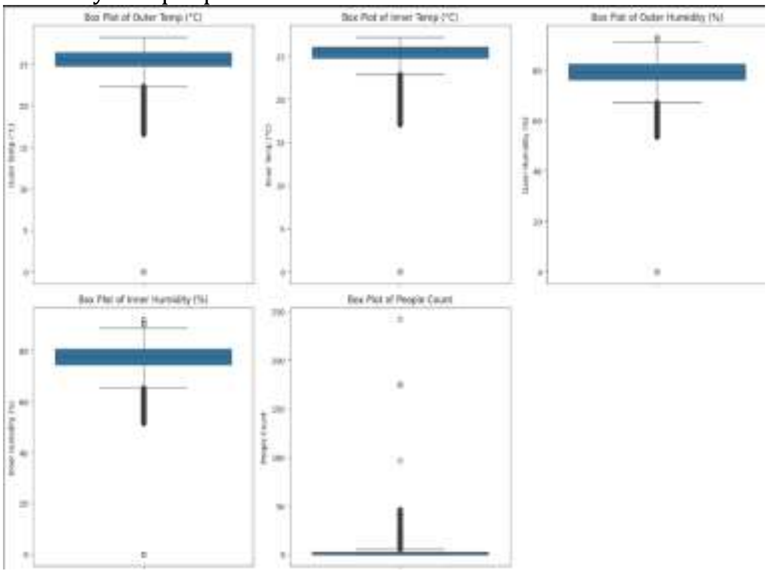

Fig. 4. Outliers detected in the recorded data

Each box plot provides a visual summary of the distribution of a numeric feature in the dataset. The key components of the plots are:

3.1 Median (Central Line in the Box)
The horizontal line inside each box represents the median (50th percentile) of the data for that feature. It indicates the central value around which most data points lie. For temperature box blot i.e. outer and inner temperature readings, most of the data lies between 15°C to 26 °C.
Outlier detected at 0°C. For humidity data values lie between 70 to 80 %. Outlier detected at 0%. In case of people count, it lies between 0 TO 40.
3.2 Interquartile Range (IQR): In the generated box plot, Q1 represents the first quartile which is greater than 25% and less than 75% of the data. Q2 represents the second quartile i.e. half of the data. Q3 i.e. third quartile is greater than 75% of the data and must be smaller than the remaining 25 % of the data. One vertical line extending from the box represents the typical values. These are the whiskers. The box extends from the 25th percentile (Q1) to the 75th percentile (Q3). This region contains the middle 50% of the data, showing where most values are concentrated. The vertical lines extending from the box represent the range of typical values. It is called whiskers. The whiskers usually go up to 1.5 × IQR beyond Q1 and Q3.
3.3. Outliers: Any points lying beyond the whiskers are marked as outliers. These represent unusually high or low values compared to the rest of the observations. In the plots, outliers are below the whisker. Removal of outliers plays an important role.

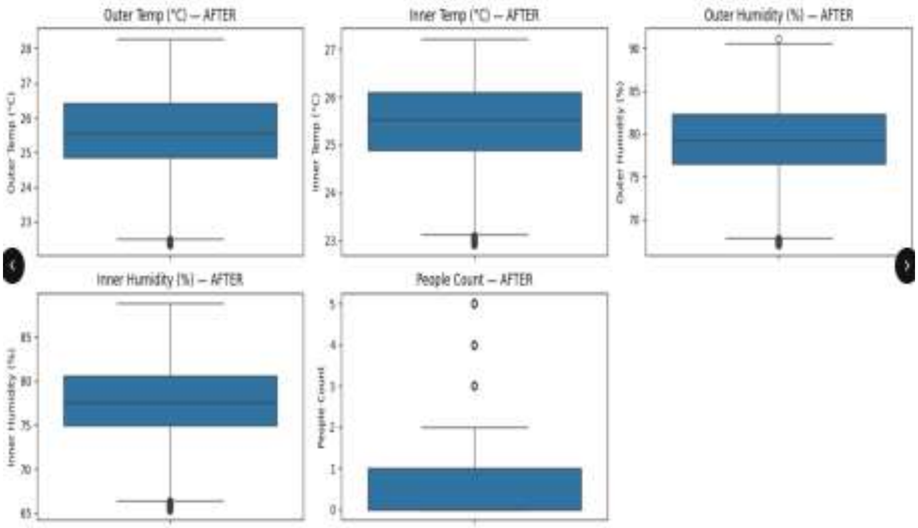

Fig. 5. Data after Outlier removal

Outlier removal is an elimination of extreme values. This step is important because it boosts data accuracy, model performance, and the reliability of statistical analyses by eliminating extreme values that can distort results. Removing outliers prevents them from skewing the mean, which leads to more robust analysis, better-performing machine learning models, and more accurate conclusions. Fig. 5 shows the box plots showing data after outlier removal.
The logistic regression model is well suited for the collected data. As the collected data mainly contains the result in the form of whether the AC will be on or off. For the binary data, the logistic regression algorithm is best suited. They are less complex as compared to other machine learning models.

To evaluate the performance of a logistic regression model, a confusion matrix can be calculated. It helps in identifying where the model makes errors by categorizing predictions into false positives, true negatives, false negatives and true positives.

Among 8122 recorded data values

- True Negatives: The model correctly predicted 5596 instances where the AC should remain OFF
- True Positives: It also correctly identified 1123 instances where the AC should be ON.
- False Positive: in 165 cases, the model incorrectly turned the AC ON when it should have stayed OFF (False Positives).
- False Negative: In 1238 cases, the system failed to turn the AC ON when required (False Negatives).

Table 1. Logistic regression model accuracy

Model	Training Accuracy	Testing Accuracy	Precision	Recall	F1-score	Training Time (s)	Training Dataset Count	Testing Dataset Count
Logistic Regression	82.97%	82.73%	0.8719	0.4756	0.6155	0.1379	32485	8122

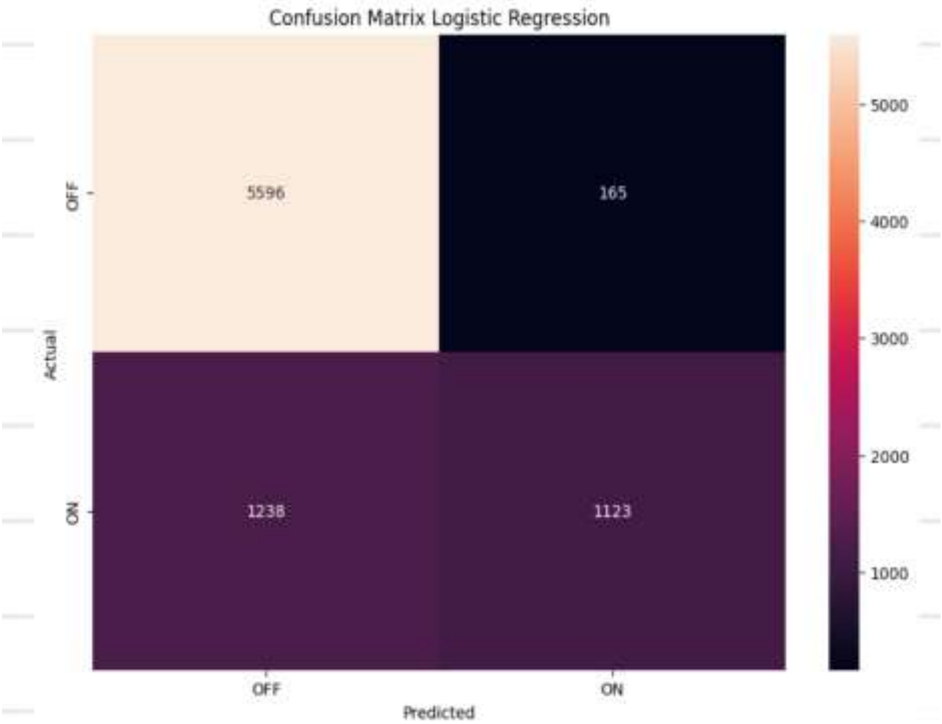

Fig. 6. Confusion matrix Logistic regression

Let us calculate the accuracy of the model -

$$\text{Accuracy} = \frac{\text{True negative} + \text{True positive}}{\text{False positive} + \text{True negative} + \text{False Negative} + \text{True positive}}$$

$$= 6719 / 8122$$

Accuracy = 82.72 %

To indicate how many of the instances predicted as "ON" were actually ON, precision value is useful. A high precision value means the model makes very few false "ON" predictions. How well the model identifies all situations where the AC should be ON, recall value is calculated. In Table 1. Shows the model accuracy. It was evaluated as 82.72% with precision and recall as 87 % and 47 % respectively. The prototype was tested in a medium-sized room accommodating up to six persons. The experimental results demonstrated that the system successfully detected occupancy with an accuracy of up to 80%, confirming the reliability of proximity-based detection. The logistic regression model achieved consistent predictions regarding the cooling requirement, resulting in improved energy efficiency and user comfort. The system effectively reduced unnecessary AC operation when the room was vacant, saving a significant amount of electricity over multiple test sessions. When occupancy increased, the automatic adjustment of cooling provided an optimal temperature environment without manual intervention. The integration of multiple sensors allowed for accurate environmental monitoring, while the ESP32 efficiently processed real-time data. Moreover, the IoT-based design ensures scalability and affordability, making it feasible for use in domestic environments, conference halls, classrooms, and spaces frequented by senior citizens.

3. CONCLUSION

This model totally runs on a remote-less platform. Hence, this model is proven to be most useful for senior citizens, at conference halls of educational institutes. Also, this model does not have the overhead of replacing non-smart AC by smart AC. This IoT model can be immersed in existing non-smart ACs. Overall, the results indicate that the proposed IoT-enabled model provides a cost-effective and intelligent solution for enhancing the functionality of non-smart air conditioners. It bridges the gap between comfort, automation, and energy conservation in an economically sustainable manner.

4. ACKNOWLEDGEMENT

I am grateful to the Computer Science Department, Vishwakarma University, Pune for providing the necessary facilities and conducive research environment that made this project possible. My thanks extend to Dr. Pooja Kulkarni and lab members who offered valuable suggestions and technical assistance during the implementation and testing phases.

REFERENCES

[1] Mohamed, K., El Shenawy, A.: A Smart IoT-Based Home Automation System for Controlling and Monitoring Home Appliances. International Review of Automatic Control.16,228–237(2023). https://doi.org/10.15866/ireaco.v16i5.23872
[2] Mena, A.R., Ceballos, H.G., Alvarado-Uribe, J.: Measuring Indoor Occupancy through Environmental Sensors: A Systematic Review on Sensor Deployment, (2022)
[3] Abdul Lateef Haroon, P.S., Shafiulla, M., Naveed, S.M., Ahmed, S., Nawaz, S.M., Kumar, U.: Home Automation Using Wi-Fi: ESP32-Based System for Remote Control and Environmental Monitoring. In: 3rd IEEE International Conference on Distributed Computing and Electrical Circuits and Electronics, ICDCECE 2024. Institute of Electrical and Electronics Engineers Inc. (2024)
[4] Ullo, S.L., Sinha, G.R.: Advances in smart environment monitoring systems using iot and sensors, (2020)
[5] Budijono, S., Felita: Smart Temperature Monitoring System Using ESP32 and DS18B20. In: IOP Conference Series: Earth and Environmental Science. IOP Publishing Ltd (2021)
[6] Tsvetanov, F.A.: Storing data from sensor networks. In: IOP Conference Series: Materials Science and Engineering. IOP Publishing Ltd (2021)
[7] Marek Babiuch: Using the ESP32 Microcontroller for Data Processing. Institute of Electrical and Electronics Engineers (2019)
[8] Roberto Pashic, PhD. ESPRESSIF ESP32 DEVELOPMENT BOARD IN WIFI STATION COMMUNICATION MODE. TEMEL-ij. 4, 1–6 (2020)
[9] Xin Zhou: Comparison of different machine learning algorithms for predicting air-conditioning operating behaviour in open-plan offices Energy and Buildings, (2021), 111347, 251
[10] R. Raghavendra: Sensor-based human activity detection using a novel machine learning algorithm in home automation Multidisciplinary Science Journal, (2024)
[11] Human Presence Detection by monitoring the indoor $CO2$ concentration Sebastian Wilhelm ACM International Conference Proceeding Series, (2020), 199-203
[12] Sathyanarayanan S: Confusion Matrix-Based Performance Evaluation Metrics. African Journal of Biomedical Research, (2024), 4023-4031
[13] Heydarian, Mohammadreza: MLCM: Multi-Label Confusion IEEE Access, (2022), 19083-19095, 10