



Smart Gas Monitoring And Prediction System Using IOT And ML

M.Kumar¹, K. Mariya Priyadarsini²

Final Year Student, Department of Electronics and Communication Engineering,
Andhra Loyola Institute of Engineering and Technology¹

Assistant Professor, Department of Electronics and Communication Engineering,
Andhra Loyola Institute of Engineering and Technology²

Abstract: The dangerous effects of various forms of air pollution - as well as the presence of carbon monoxide (CO), ammonia (NH₃), carbon dioxide (CO₂), and other hazardous gases - pose a significant threat to both human health and the safety of our environment, especially in indoor spaces and in industrial settings. Because there are very few continuous monitoring systems, there is often a delay in detecting hazardous gaseous conditions, which increases the likelihood of accidents and negatively impacts health. This research project describes the design of a smart gas monitoring and prediction system that combines the Internet of Things (IoT) with Machine Learning (ML) to allow for the real-time monitoring and prediction of the concentrations of hazardous gases. This system is constructed using an ESP32 microcontroller, integrated with MQ-7, MQ-137, and SCD40 sensors to monitor CO, NH₃, CO₂, temperature, and humidity. All sensor data is captured and streamed to a ThingSpeak cloud service for storage, analysis, and visualization. To perform predictive analysis, a hybrid deep learning model consisting of Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks is constructed to analyze the time series data and forecast future levels of hazardous gases. This hybrid deep learning model is intended to enhance the accuracy of hazard detection and provide an earlier indication of dangerous conditions. A web-based dashboard has also been developed, which provides both the real-time and predicted values of the sensor measurements. Ultimately, the smart gas monitoring and prediction system will provide continuous monitoring of gaseous conditions, provide notifications of any hazardous conditions in a timely manner, and provide the opportunity to take action by implementing proactive safety measures; thereby improving the overall quality of environmental monitoring and increasing safety for the public.

Keywords: Internet of Things (IoT), Smart Gas Monitoring, Machine Learning (ML), Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), ESP32, MQ-7 Sensor, MQ-137 Sensor, SCD40 Sensor, Carbon Monoxide (CO), Ammonia (NH₃), Carbon Dioxide (CO₂), ThingSpeak, Air Quality Monitoring, Predictive Analytics

I. INTRODUCTION

Recently, air pollution has become a major global problem due to its impact on people's health and the planet's health. Many air pollutants, such as ammonia (NH₃), carbon monoxide (CO), and carbon dioxide (CO₂), can cause serious health problems like respiratory illnesses, poisoning, and even death if they are allowed to escape without detection. In addition, gas leaks and the degradation of air quality can go unnoticed in some cases; as a result, responses are delayed, and the potential for harm is increased.

While advances in technology, such as IoT (Internet of Things) and ML (Machine Learning), create new opportunities for real-time monitoring of environmental conditions and predicting future outcomes, they also allow for the collection of continuous data from sensors. By using these technologies together, organizations can identify problems before they occur and have more time to make safe choices.

This project involves utilizing ESP32 microcontroller along with multiple sensors, such as MQ-7, MQ-137, and SCD40 to develop a Smart Gas Monitoring and Prediction System that detects gases as well as temperature and humidity, and sends the information to ThingSpeak which allows for the collection, storage and visualisation of the data. The hybrid deep learning model, LSTM and GRU will be used to process the time series dataset and predict future gas levels based on the previous data.

This project aims to provide continuous monitoring, early warning alarms, and ultimately increase safety by providing the user with sufficient time to take action before a dangerous situation occurs. This research will contribute to improving environmental monitoring as well as providing safer living and working environments.



II. LITERATURE SURVEY

Air quality monitoring has become a critical research area due to the severe health and environmental risks caused by hazardous gas emissions. In many developing and rural regions, communities are exposed to gas leaks originating from agricultural activities, small-scale industries, and local factories, which often remain undetected until serious health effects occur due to the absence of continuous monitoring systems in remote locations [1]. Conventional air quality monitoring approaches rely on centralized stations and manual inspection methods, which are effective in urban areas but costly and impractical for rural deployment. Limited infrastructure, high maintenance costs, and poor connectivity further prevent real-time data collection and analysis in geographically isolated regions, resulting in delayed responses from healthcare providers and authorities only after exposure has caused harm [2]. Recent advancements in Internet of Things (IoT) technology have enabled low-cost, distributed sensor networks capable of real-time environmental data collection. In parallel, Machine Learning (ML) techniques, particularly time-series models such as Long Short-Term Memory (LSTM) networks, have demonstrated strong performance in analyzing and predicting air quality trends [3]. According to the World Health Organization (WHO), air pollution remains one of the most significant global public health threats, with nearly nine out of ten people worldwide exposed to polluted air, leading to increased mortality and severe health complications [4]. The WHO Handbook on Indoor Radon identifies indoor gas exposure as a critical concern, emphasizing the importance of continuous monitoring and preventive measures to reduce long-term risks [5]. Furthermore, Almendra et al. [6] highlighted the relationship between social deprivation and excess mortality, showing that environmentally and economically vulnerable regions experience higher health risks, thereby reinforcing the need for real-time monitoring systems and targeted intervention strategies.

III. ALGORITHM

The Smart Gas Monitoring and Prediction System is designed according to an algorithm with real-time data collection via sensors, computation on the cloud, and prediction using machine learning. The following is a breakdown of the complete workflow of the system:

3.1 Initialize the Sensors

Power on the ESP32 microcontroller and connect it to the three gas sensors MQ-7 (Carbon Monoxide), MQ-137 (Ammonia), and SCD40 (Carbon Dioxide, Temperature and Humidity) sensors; make sure that all sensors have been calibrated in order to have an accurate reading.

3.2 Acquire Data

Once the sensors have been initialized, and are continuously collecting real-time environmental data from all sensors at set intervals; the ESP32 will be processing both the analog and digital signals being produced from all sensors.

3.3 Transmit Data to the Cloud

All sensor data will be sent to the ThingSpeak cloud using Wi-Fi; this allows for storage of the data collected and visualisation of that data.

3.4 Data Preprocessing

Retrieve and pre-process the stored data. "Pre-processing" includes clean-up of data (to eliminate errors), normalisation (to bring values to equivalent scales) and deal with missing/zero value data sets, which will enhance the prediction accuracy of the model that will be created from this data.

3.5 Model Training

Create a hybridised LSTM/GRU deep learning model and train it to learn and predict the gas concentration values from past historical time-series.

3.6 Prediction

Once the model has been created and trained, it will predict gas concentrations based on current input data.

3.7 Visualisation

Perform real-time web-based display of both current gas concentrations and future predicted gas concentrations.

3.8 Alert generation

If the gas concentrations exceed the preset thresholds, trigger an alert (e.g., noise notifications, system alerts, etc.) to indicate potential safety hazards to personnel on the site.



3.9 Continuous Monitoring

The process runs continuously throughout an uninterrupted loop, ensuring continuous monitoring, predicting, and alerting. The system takes advantage of periodic updates from new data sources to enhance the training of models over time.

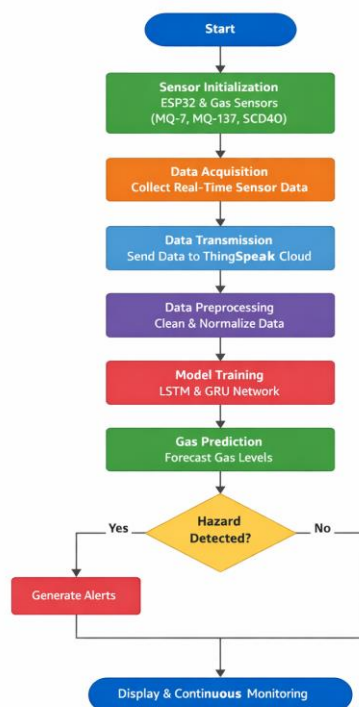


Figure 1: Flowchart of Smart Gas Monitoring and Prediction System

IV. METHODOLOGY

Using an Internet of Things (IoT) approach method along with Machine Learning (ML) techniques, we will have the Smart Gas Monitoring - Prediction System. It will let us monitor, analyze, and predict potentially harmful gas levels in real time. The proposed system will use the layered and modular approach to provide an accurate representation of gas levels for each of the following: sensor layer, communication layer, data processing layer, prediction layer, and alert generating layer. This configuration provides the following benefits: dependable, scalable, efficient, and effective, regardless of your indoor or industrial environment.

4.1 Data Acquisition

Data acquisition takes place through the use of multi gas sensors interfaced with an ESP32 microcontroller. To detect CO sensors we are using MQ-7, for NH₃ we are using MQ-137 and for CO₂ along with temperature & humidity we are utilizing SCD40. Collectively these sensors will provide an approximate measurement of CO, NH₃, CO₂, temperature, and humidity. All of these sensors provide analog or digital signals that are proportional to their respected gas concentration level. We also quantify the sensors' readings from an analog sign by utilizing the ESP32's internal ADC and convert these values into calibrated numbers for all parameters mentioned above. The sensors need to be calibrated after they have been manufactured, in order to create an accurate measurement and to reduce environmental factors. Data will be recorded at particular intervals based on studies involving time series.

4.2 Data Transmission & Integration with Cloud

Data that has been processed will be sent to the ThingSpeak cloud via Wi-Fi through the ESP32. Once the data has been uploaded to ThingSpeak it can be securely stored, visually represented and analyzed immediately. Each parameter, CO, NH₃, CO₂, temperature and humidity, that was collected from the sensors is uploaded to the ThingSpeak cloud as a different field within the database. The ThingSpeak platform allows for API-based access to the stored parameters, allowing for easy and direct access for use with machine learning based models and web dashboards. This cloud-based architecture allows for remote access of all stored data, reducing the need for local data storage.



4.3 Data Cleaning

Raw sensor readings often contain noise, have missing data, or may vary because of environmental factors or instability in the network environment. Therefore, preprocessing techniques like filtering out noise, deleting outliers from the dataset, normalizing the data to a standard range to promote rapid convergence of the model during training, and using interpolation (intermediate values between two points) to fill in missing data points must be carried out. Once the sensors have been cleaned; the final version of the sensor data is placed into a format suitable for time series analysis. Since many deep learning networks focus on time series and sequential data patterns, this last preprocessing step is very important in order for the model to learn correctly.

4.4 Feature Engineering and Creation of Time Series

During this step, features relevant to the application or model are selected and organized in sequence to define the temporal relationship between them. In order for models trained with past observation(s) to be able to accurately predict the future value, the training of the model requires generating input-output (IO) pairs using a sliding window method. The use of this method allows the model to identify trends, seasonality, and variability in gas concentration over time. Using feature scaling and creating a sequence of sensor data with the above methods will greatly improve the performance of your model.

4.5 ML Model (LSTM + GRU)

The system employs a hybrid deep learning architecture with Long Short-Term Memory (LSTM) networks and Gated Recurrently Unit (GRU) Networks. The LSTM layers are able to model long-term dependencies and solve the vanishing gradient problem while the GRU layers train faster and have lower computational complexity. The hybrid model implements both LSTM and GRU networks in a way that uses their strengths to provide better predictive accuracy. The model is trained with historical sensor data using hyperparameters that have been optimized for the application, such as learning rates, batch sizes, and number of epochs. The loss function used during training to evaluate the performance of the models is the Mean Squared Error (MSE).

4.6 Prediction and Decision Support

After training, the model is deployed to predict the future concentration level of gas based on the incoming real-time data. Continuous predictions are generated and compared with pre-established safety thresholds which have been set for each of the gases and were based on established safety regulations. If predictions exceed established safe limits, the system identifies a potentially hazardous condition based on its predictive capability and therefore detects potential risks before they reach a critical level.

4.7 Alert Generation and Notification System

The system serves a purpose of notifying the user with respect to possibly hazardous conditions. As a result, the system generates alerts via buzzer sound, visual warnings on the dashboard, as well as cloud notifications. Alert system design was undertaken to generate few false alarms while yielding high sensitivity to hazardous conditions.

4.8 Data Visualization

The system provides a web-based dashboard that displays real-time (actual) sensor readings and a tallied predicted value for each sensor at a given time. The web-based dashboard serves as an "intuitive" interface for monitoring environmental conditions, trend analysis, and historical (and real-time) data analysis. The use of visualization tools such as line graphs and charts improves understanding, therefore encourages decision making.

4.9 Continuous Monitoring of Sensors and Model Updating

The system continues to operate in a loop to provide ongoing (non-interrupted) monitoring and prediction of hazard conditions. The data that is continuously captured from the sensors is used to periodically retrain (or update) the machine learning model and improve its accuracy. This continuous learning ability allows the system to quickly adapt to variations in environmental conditions.

V. ARCHITECTURE

5.1 System Architecture Overview

This smart gas monitoring and prediction system is based on an IoT, cloud computing and machine learning architectural framework whereby each individual component of these three technologies interact together. The smart gas monitoring and prediction system allow for real-time monitoring, analysis and predictions of potentially dangerous levels of gas in both indoor and industrial settings.



The architecture of the smart gas monitoring and prediction system is composed of many different layers which come together to provide the best possible means for collecting, processing, predicting and generating alerts for gas. Each layer of the smart gas monitoring and prediction system has its own purpose but interfaces seamlessly with each of the other layers.

The sensing layer gathers environmental data (air) through the use of gas sensors; the communication layer is responsible for transferring that data reliably to the cloud platform; the cloud/data processing layer provides data storage/pre-processing/management of the data, which is then available for future analysis; and the machine learning layer will use a hybrid LSTM and GRU machine learning model to perform predictive analysis based on the historical pattern of previous gas concentration levels, in order to predict the future concentration of gas(s)(in combination with the first three layers).

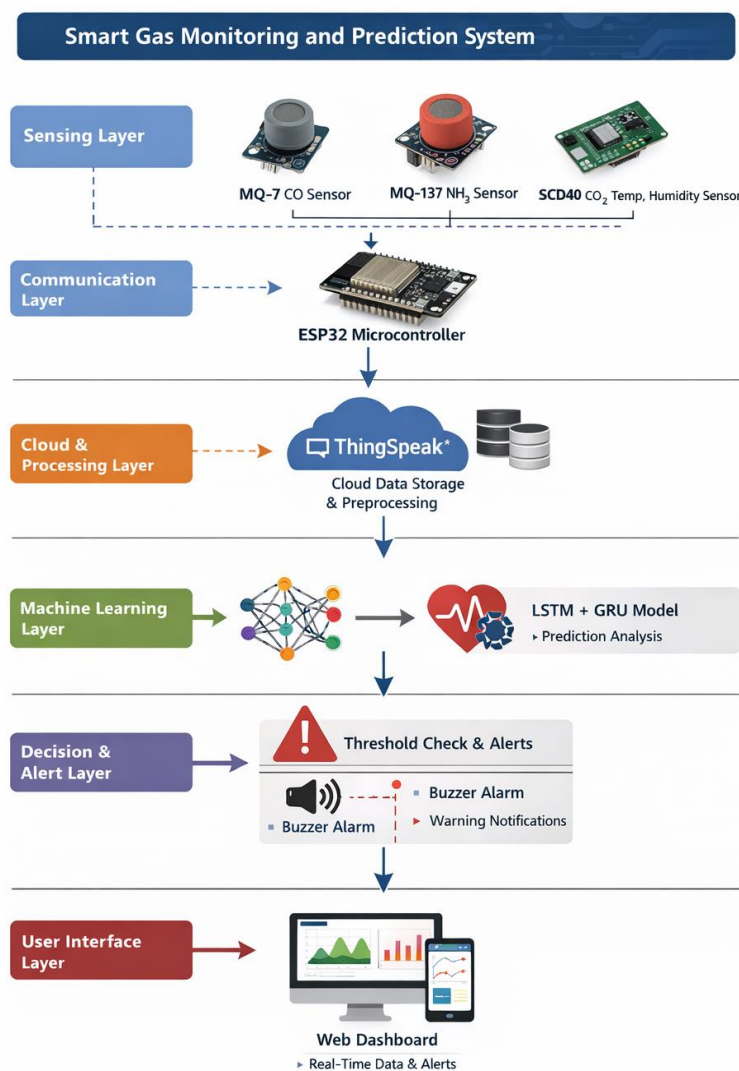


Figure 2: System Architecture of Smart Gas Monitoring and Prediction System

5.2 Architecture Layers

1. Sensing Layer (IoT Layer)

The sensing layer of the overall architecture of the Internet of Things (IoT) consists of several types of gas/chemical sensors and other environmental conditions (temperature, humidity). Examples of these sensors include MQ-7, MQ-137, and SCD40 gas/chemical sensors, which are connected to an ESP32 microcontroller using the ESP32 microcontroller's analog and digital I/O pins. The gas/chemical sensors continually monitor CO (carbon monoxide), NH₃ (ammonia), CO₂ (carbon dioxide), etc. These sensors send their data to the ESP32, which collects and processes the data using an on-chip



Analog (ADC), preparing it for transmission as digital signals to the next layer (Communication Layer) of the architecture. This layer is the first major layer required for the system to function.

2. Communication Layer

The communication layer is responsible for transmitting sensor data from the ESP32 to a cloud platform (ThingSpeak). The ESP32 uses Wi-Fi to connect to the internet and sends the data to the ThingSpeak server via HTTP protocols, providing a reliable and real-time transmission of the sensor data to the cloud, even in areas with moderate wireless connectivity.

3. Cloud and Data Processing Layer

The ThingSpeak cloud platform assists in receiving newly received sensor data and can prepare the incoming data for machine learning analysis. Furthermore, this layer provides an application program interface (API) to assist with visualizing the data as well as performing some of the preprocessing tasks on the data (e.g., data cleaning, normalization, formatting), which can be carried out either in the cloud or in conjunction with an external processing environment connected to the cloud (e.g., Python).

4. Machine Learning and Prediction Layer

In this layer, a hybrid LSTM and GRU model has been developed to allow for time-series prediction from historical sensor data. The algorithm will be trained with prior sensor data in order to make predictions of what future gas concentrations will be. Additionally, the algorithm will analyze temporal patterns and trends so that accurate future predictions of gas concentrations can be made based on historical sensor data collected in prior periods. This predictive capability is what represents the intelligence of the overall system.

5. Decisions on Alerts and Decision Support Layer

In this layer, gas value prediction are compared to predefined safety thresholds. If the predicted gas levels exceed the safe limits, alerts will be triggered in the form of activated buzzers and warning notifications. This layer provides timely decision-making ability to help avoid hazardous situations.

6. User Interface Layer

The final layer provides users with a web-based dashboard for real-time sensor data display, prediction values, and historical trends; all of which will be in graphical or chart form. Users will be able to remotely access the web-based dashboard and monitor environmental conditions and take any needed action based on alerts.

VI. RESULTS AND ANALYSIS

To evaluate the proposed Smart Gas Monitoring and Prediction System, actual sensor data collected from MQ-7, MQ-137, and SCD40 is used. The database consists of time series values of gas concentration in addition to time series values of temperature and humidity. The time series values of gas concentrations, temperature, and humidity will be used to train and test both the Hybrid LSTM and the GRU model. The evaluation focuses on predicting the correct outcomes and on detecting potential threats/hazards.

6.1 Evaluation Metrics

Standard regression metrics including the Root Mean Square Error (RMSE) and R^2 score have been used in order to evaluate the performance of this model. These metrics compare the actual value and the predicted gas concentration values.

Root Mean Square Error (RMSE) is the error associated with the prediction of a value and is measured using the same units as the data.

6.2 Experimental Results

The R^2 score indicates the degree to which the model represents the dataset; the closer to 1 the better model performance. The hybrid model of LSTM and GRU models was used for training and validating the predictive performance of the model based on historical data through the combination of LSTM and GRU. Figure 1 shows the results obtained from testing the hybrid LSTM-GRU model. The RMSE is 6.9 and the R^2 Value is 0.85. Based on these two statistics, the model provides high rates of accuracy in prediction with only a very small amount of error made per prediction. The RMSE value indicates there is very little difference between the actual values versus those predicted by the model, and the R^2 Value indicates that there is significant correlation between the actual values and those predicted by the hybrid LSTM-GRU model.



6.3 Performance Analysis

The hybrid model identified was able to identify trends within time series sensors through temporal dependencies. The predictive capabilities of the proposed approach compared to historical thresholds provide valuable foresight into potentially hazardous conditions.

The ability of the system to learn patterns and trends in gas concentrations is confirmed by the higher R^2 values. The addition of GRUs with LSTMs improved prediction accuracy as well as reduced the time taken to compute predictions. This allowed the system to predict potential hazards before they occurred, which allowed for timely preventative actions.

6.4 Visualization & Description

The actual versus predicted value comparisons visually confirm how closely related they are. This demonstrates how successful the model is. In addition, the real-time dashboard displays actual sensor data along with predicted trends for real-time monitoring of environmental conditions by users.

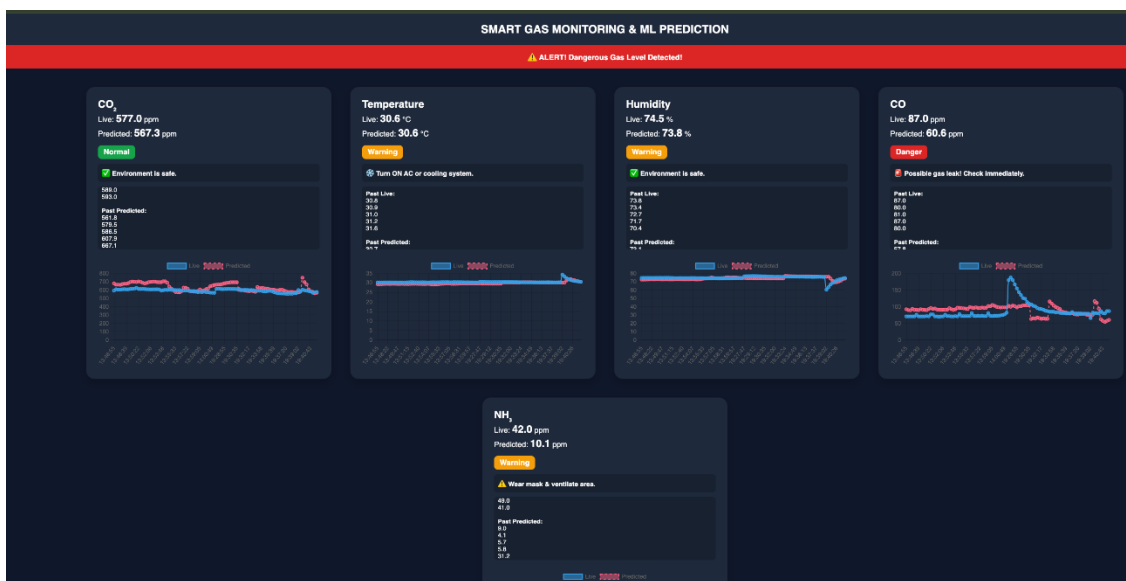


Figure 3: Web-based Dashboard for Real-time Gas Monitoring and Prediction

VII. CONCLUSION

In this work, we have created a Smart Gas Monitoring and Forecasting System using IoT and ML. This system allows real-time monitoring of air quality and prediction of hazards by collecting continuous data on hazardous gases (CO, NH₃ and CO₂) and environmental conditions (temp. & humidity) through the use of an ESP32 microcontroller and various types of sensors (MQ7, MQ137 and SCD40). By integrating these components together, we enable accurate and continuous data collection in both indoor and industrial environments.

The data that is collected from these sensors will be sent to the ThingSpeak cloud platform where it can be stored, processed and visualized. To model the time-series data for gas concentration prediction, we created a hybrid deep learning model that combines LSTM and GRU neural networks. The results of our experiments show that this model has produced a high degree of accuracy in predicting future gas concentrations, as evidenced by an RMSE of 6.9 and an R^2 of .85 providing a strong correlation to actual values and thereby providing reliable predictions.

Our innovative monitoring solution provides both real-time data and a predictive capability that allows for early detection of hazardous conditions. The integrated alerting mechanism provides for numerous methods (buzzer or warning light on the dashboard) whereby quick action can be taken to save human lives and prevent property loss.

Our system is far superior to traditional monitoring systems in that it provides greater effectiveness, less response time, and allows for proactive decisions to be made by the user.



In addition, the system is also highly cost-effective, will work on a modular basis, and has a range of uses. Applications can be found in the home, industry, and laboratory settings. The modular system will be able to accommodate additional sensors and sophisticated analytic techniques as the need arises.

There are several areas for continued research and development, including improving accuracy of predictions through the use of advanced deep learning models, implementing edge computing so that data can be processed and transmitted in real-time, and expanding the overall capabilities of the system to monitor a larger environment.

The proposed monitoring system is a major enhancement for public safety, environmental monitoring, and the development of IoT-based smart solutions.

REFERENCES

- [1]. W. Hu, Y. Wen, K. Guan, G. Jin, and K. Tseng, "iTCM: Toward learning-based thermal comfort modeling via pervasive sensing for smart buildings," *IEEE Internet of Things Journal*, vol. 5, no. 5, pp. 4164–4177, 2018.
- [2]. H. Li, S. Li, J. Yu, Y. Han, and A. Dong, "AIoT platform design based on front and rear end separation architecture for smart agriculture," in *Proc. 4th Asia Pacific Information Technology Conf.*, 2022, pp. 208–214, doi: 10.1145/3512353.3512384.
- [3]. B. Silva, J. Rodrigues, A. Ramos, K. Saleem, I. Torre, and R. Rabelo, "A mobile health system to empower healthcare services in remote regions," in *Proc. IEEE Int. Conf. on E-Health Networking (HealthCom)*, 2019, pp. 1–6.
- [4]. World Health Organization, "9 out of 10 people worldwide breathe polluted air, but more countries are taking action," WHO, 2018. [Online]. Available: <https://www.who.int/news-room/detail/02-05-2018-9-out-of-10-people-worldwide-breathe-polluted-air-but-more-countries-are-taking-action>
- [5]. *World Health Organization, WHO Handbook on Indoor Radon: A Public Health Perspective, Geneva, Switzerland: WHO, 2009.*
- [6]. A. Almendra, et al., "Association between social deprivation and excess winter mortality in Portugal," *International Journal of Biometeorology*.
- [7]. L. P. Rongali, "The trinity of cybersecurity: DevSecOps, cloud security and cryptography in the digital age," *SSRN Electronic Journal*, 2025. doi: 10.2139/ssrn.5229585.
- [8]. L. P. Rongali, "Integrating AI and DevOps practices to develop cybersecurity frameworks that enhance resilience in utility infrastructure," *Journal of Informatics Education and Research*, vol. 5, no. 2, 2025. doi: 10.52783/jier.v5i2.2838.
- [9]. A. Todupunuri, "Develop machine learning models to predict customer lifetime value for banking customers," *International Journal of All Research Education & Scientific Methods*, vol. 12, no. 10, pp. 1254–1259, 2024. doi: 10.56025/ijaresm.2024.1210241254.
- [10]. A. Todupunuri, "Explore how AI can be used to create dynamic and adaptive fraud detection systems in digital banking," *SSRN Electronic Journal*, 2024. doi: 10.2139/ssrn.5014699.