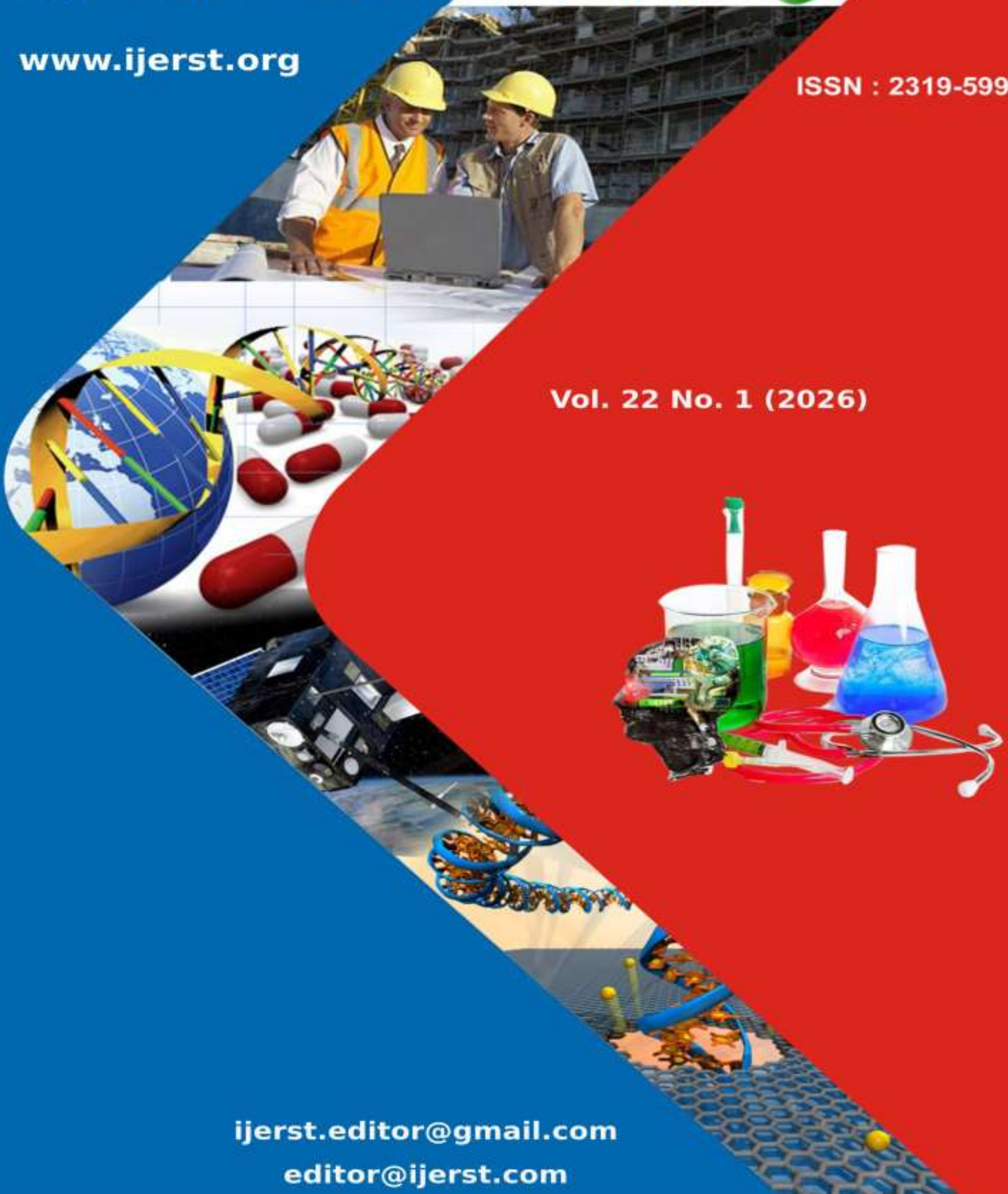
IJERST

International Journal of Engineering Research and Science & Technology

www.ijerst.org

ISSN : 2319-599



Vol. 22 No. 1 (2026)

ijerst.editor@gmail.com

editor@ijerst.com

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com
Vol. 22, No. 1, 2026
© 2026 IJERST. All Rights Reserved
Peer Reviewed, Referred & Indexed Journal

Research Paper

A Machine Learning-Based Platform for Monitoring and Prediction of Hazardous Gases in Rural and Remote Areas

Mrs.K.Bharathi, M.Tech - Assistant Professor, Department of MCA, Bapatla Engineering College, Bapatla, Andhra Pradesh
Ms. Shaik Mahee Jabeen, Reg No: Y25MC23078, Ms. Kalli Sirisha, Reg No: Y25MC23029.
Ms. Boddu Hema Latha, Reg No: Y25MC23012, Mr. Redithi Hemanth Brahmam, Reg No: Y25MC23069,
Department of MCA, Bapatla Engineering college, Bapatla, Andhra Pradesh

***Abstract*—Air pollution and the release of hazardous and radioactive gases pose serious threats to the environment and human health, often leading to environmental disasters and premature deaths. These risks are especially severe in remote and rural areas, where access to healthcare, emergency services, and continuous monitoring is limited. In such regions, the absence of real-time monitoring systems makes it difficult to detect indoor gas-related incidents early, causing authorities and healthcare providers to respond only after serious harm has occurred.**

To address this challenge, this paper presents a digital decision support system that combines Internet of Things (IoT) technology with Machine Learning (ML) to monitor and predict indoor hazardous gas incidents. The system is built on the Rural THINGS IoT platform, developed by the University of Beira Interior, Portugal, which continuously monitors air quality in rural and remote environments. By collecting real-time sensor data, the platform helps assess environmental conditions and warns residents and stakeholders about potential exposure risks.

The proposed system uses an advanced ML architecture that combines bidirectional and unidirectional Long Short-Term Memory (LSTM) layers to analyze time-series data and predict future gas concentration trends. Validation using a real testbed demonstrated that the model effectively predicts hazardous gas patterns, enabling timely alerts and preventive actions. This approach supports improved public safety, proactive healthcare responses, and better environmental risk management in underserved regions.

Keywords:1. Internet of Things (IoT), 2. Machine Learning (ML), 3. Long Short-Term Memory (LSTM), 4. Indoor Air Quality Monitoring, 5. Public Health.

I. INTRODUCTION

The bad effects of radioactive gases on the air we breathe are a big problem for people’s health and the environment. This is a concern all over the world. Air pollution from these gases can cause health problems, breathing issues and even death if people are around them for too long.

In cities we have systems to respond to emergencies and equipment to monitor the air, which is really helpful. But in areas and places that are far away from cities people are still very much at risk, from poisonous and radioactive gases. Early detection of indoor gas-related incidents is very difficult in these areas because they typically lack real-time alerts, ongoing environmental monitoring, and rapid access to healthcare and safety services.

In the country, where people live in villages and there are lots of farms and little factories bad gas leaks can happen and nobody notices until something really bad has already happened. The old ways of checking for these leaks are usually too costly or just not practical to use in places that’re really far away from cities. So the people, in charge and the doctors usually only do something after the gas has already hurt someone. This shows that we really need to find a way to watch the air quality all the time that’s smart does not cost a lot and can be used in many places. We need this to happen so we can know when something bad is going to happen with the air before it becomes very dangerous. Air quality is a deal and we need to monitor air quality to keep people safe.

To address this problem, this paper presents a Machine Learning-based digital platform built on an Internet of Things (IoT) ecosystem for real-time indoor monitoring of environmental gases. By combining sensor data with advanced time-series prediction models, the platform not only detects current air quality conditions but also forecasts future gas concentration trends. This predictive capability enables early warnings, timely interventions, and improved protection of public health in rural and remote communities.

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. These incidents often remain undetected until serious health effects occur, mainly 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 prevent real-time data collection and analysis in geographically isolated regions. As a result, healthcare providers and authorities typically respond only after exposure has already caused harm, highlighting the need for proactive monitoring solutions [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. The WHO report on global exposure to air pollution states that nearly nine out of ten people worldwide are exposed to polluted air, leading to increased mortality and severe health complications [4]. The report emphasizes that low-income and rural populations are disproportionately affected due to inadequate monitoring systems and limited policy enforcement, reinforcing the need for real-time air quality monitoring and early warning mechanisms. The WHO Handbook on Indoor Radon [5] identifies indoor radon exposure as a critical public health concern, particularly due to its strong association with lung cancer. The handbook outlines radon sources, health impacts, measurement methods, and mitigation strategies. It stresses the importance of continuous indoor gas monitoring and preventive policies to reduce long-term exposure risks.Almendra et al. [6] examined the relationship between social deprivation and excess winter mortality in Portugal. Their spatial analysis revealed higher mortality rates in socioeconomically deprived regions, influenced by environmental conditions, housing quality, and access to healthcare. The study highlights regional inequalities in environmental health risks and supports the need for targeted monitoring and intervention strategies in vulnerable communities.

III. ALGORITHM

Hazardous Gas Monitoring and Prediction

Input: Real-time sensor data (gas concentration, tempera- ture, humidity, timestamp)

Output: Predicted gas concentration levels and alert notifi- cations

Steps:

1. Deploy IoT gas sensors in indoor environments to continuously collect air quality data.

2. Transmit sensor data to the cloud platform using wireless communication protocols.

3. Preprocess the collected data by removing noise, handling missing values, and normalizing features.

4. Organize the data into time-series sequences suitable for model training and prediction.

5. Train a stacked Long Short-Term Memory (LSTM) model using historical sensor data.

6. Use the trained model to predict future gas concentration levels.

7. Compare predicted values with predefined safety thresholds.

8. Generate alerts and notifications if predicted levels exceed safe limits.

9. Store predictions and historical data for visualization and further analysis.

IV. METHODOLOGY

The Internet of Things technology is used with Machine Learning techniques to monitor and predict gases in rural and remote areas in real time.

The system has a structure that includes

* collecting data
* processing the data
* making predictions, with the data
* sending out alerts when something is wrong.

The Internet of Things and Machine Learning are used together to make this work.

The system is made to watch for gases and send alerts in real time.

A. Data Acquisition

Gas sensors that do not cost a lot are used in homes in the country small factories and farms. These gas sensors always check the air for gases and also check the temperature and how humid it is. The information that the gas sensors collect is sent away to a computer system, on the internet using special radios that can send information through the air.

B. Data Preprocessing

The raw sensor data that we get can have some problems. It may have noise in it. Some values might be missing. Sometimes things in the environment or issues with the network can cause these problems. So we need to do some things to the data to make it better. We do things like clean the data make sure it is all, at the level and get the time right. After we do these things to the sensor data we put it into an order. We make it into time-series sequences so that it works well with Machine Learning models.

C. Machine Learning Model

To understand how gas levels change over time we use a kind of computer program called a stacked Long Short-Term Memory neural network. This program has lots of layers that help it learn from the past and the future at the time. It has some layers that look at things from both sides and some that just look one way. This helps us make guesses about what will happen next. We train this program using sensor data and we try to make it as good, as possible at predicting what will happen so it makes fewer mistakes.

D. Prediction and Decision Support

The system we have can tell us what the gas levels will be like in the future. It does this at times. These gas levels are then compared to what we think is safe. If the gas levels are going to be too high the system sends out warnings to people who live there to doctors and to the people, in charge. The system sends these warnings so that everyone knows what is going on with the gas levels.

E. System Validation

The proposed platform is validated using a real-world testbed deployed in rural environments. Performance metrics such as prediction accuracy and reliability are evaluated to demonstrate the effectiveness of the system.

V. ARCHITECTURE

Architecture Overview

The system they are talking about is made to keep an eye on air inside buildings in the countryside and other far away places all the time. It can even guess when something bad is going to happen. This system has parts that work together including internet connected sensors, big computers on the internet and special computer programs that help make decisions specifically Machine Learning-based decision support, for the system and the system is called the proposed system architecture.

1. Sensing Layer (IoT Layer)

This layer is made up of cheap gas sensors that are put in places like homes in the country, farms and small factories. The gas sensors are always checking the air for gases and they also check the temperature and how humid it is. The gas sensors are the starting point, for collecting information in time about the gas concentrations and the environment. The gas sensors are really important because they help us get the information we need from the gas sensors.

2. Communication Layer

The sensor data that we collect is sent to the cloud platform using communication technologies like Wi-Fi, GSM or LoRa. This layer makes sure that the sensor data is transferred reliably in places that are really far away, from everything and do not have very good connectivity.

3. Cloud and Data Processing Layer

The cloud layer gets a lot of sensor data and stores it. This is where the cloud layer makes sure the data is good to use. It does things like remove noise from the sensor data make sure all the sensor data is, on the scale and deal with sensor data that is missing. Then the cloud layer puts the sensor data into a format that shows what happened over time. This format is great for Machine Learning to look at the sensor data and figure things out.

4. Machine Learning and Prediction Layer

This layer uses the Machine Learning model that is based on stacked Long Short-Term Memory networks. It looks at new data to figure out what the levels of hazardous gas will be in the future. The Machine Learning model is really good, at predicting things. This means we can find out about situations before they happen. The Machine Learning model helps us do this by looking at the new data and predicting what the hazardous gas levels will be.

5. Decision Support and Alert Layer

The amount of gas in the air is checked against levels that we already know. If the gas levels are not safe or if we think they will become unsafe the system sends out warnings. These warnings are sent to people who live in the area to doctors and nurses and to the people in charge. They get these warnings on their phones or, on websites. The gas levels are always checked to make sure everyone is safe.

6. User Interface Layer

This layer provides dashboards for real-time visualization of air quality, historical trends, and prediction results. It supports informed decision-making and timely intervention.

VI. EXPERIMENTAL ANALYSIS

This part is about how we set up our experiment, what the data looks like how we measure success and how well the Machine Learning-based hazardous gas monitoring platform actually works. We did these experiments to see if the system is really good at figuring out how much hazardous gas is, in the air and sending out warnings on time in rural and remote indoor areas where the Machine Learning-based hazardous gas monitoring platform is used.

A. Experimental Setup

The equipment we are using includes Internet of Things devices that can detect gas in the air. These devices are placed in homes in areas and in small industrial spaces. Each device has sensors that can find bad gases in the air. They also have sensors that can tell us the temperature and how humid it's

The Internet of Things devices send the information they collect to a computer server at regular times. The data from the Internet of Things devices is stored in one place and used for two things: to keep an eye on things as they happen and to train models. The Internet of Things devices and the data they collect are very important, for this project.

B. Dataset Description

The dataset we used for this study was collected over a weeks from the sensors that were set up. The dataset has information about the gas concentration in the air at times and it also has details about the environment at those times. Before we started training we split the dataset into two parts: one for training and one for testing. We usually split it so that 70 percent is for training and 30 percent is for testing. We did some things to make the data better like removing values that did not make sense making sure all the values are, on the scale and dealing with missing information. This helps make sure the dataset is good and that the model works well.

C. Model Training and Configuration

The prediction model we are talking about is made up of a stacked Long Short-Term Memory architecture. This Long Short-Term Memory model uses bidirectional Long Short-Term Memory layers and then unidirectional Long Short-Term Memory layers to get a sense of the temporal dependencies in the time-series sensor data. We trained the Long Short-Term Memory model using data and we made sure the settings were just right including how fast it learned how much data it looked at each time and how many times it went through the data. The goal of training the Long Short-Term Memory model was to make good predictions and make sure it worked well with new data.

D. Evaluation Metrics

To see how well the proposed system works we used some ways to measure how good a system is at predicting things. We looked at the Mean Absolute Error, the Root Mean Square Error and how accurate the predictions were. These things help us understand how different the predicted gas concentration values are, from the values. They also help us figure out if the model is really reliable. We used these metrics to evaluate the performance of the proposed system and the gas concentration values it predicts.

E. Results and Discussion

Experimental results demonstrate that the proposed LSTM-based model accurately predicts future hazardous gas concentration trends. The predicted values closely follow actual sensor readings, indicating strong learning of temporal patterns. Compared to traditional threshold-based monitoring approaches, the proposed system provides early warnings, allowing preventive actions before gas levels reach dangerous thresholds. The results confirm that integrating IoT with Machine Learning significantly enhances monitoring efficiency and public safety in rural and remote regions.

1. Perfomance Evaluation:

The performance of the proposed Machine Learning-based platform was tested to see how well it does a few things. It was checked for prediction accuracy, which's how correct the predictions are. The reliability of the alerts was also checked, to make sure the alerts are trustworthy.. The computational efficiency of the Machine Learning-based platform was checked, to see how well it can work with the resources it has. All of this was done to see if the Machine Learning-based platform is a choice for monitoring hazardous gas in real time especially in rural and remote areas where it is really needed. The Machine Learning-based platform has to be good, at predicting and alerting people to gas and it has to be able to do it quickly and efficiently.

A. Evaluation Metrics

We looked at how our prediction did, by using some standard measures. We used Mean Absolute Error and Root Mean Square Error to see how good our prediction was. Mean Absolute Error and Root Mean Square Error are two metrics that help us understand our prediction performance. We defined Mean Absolute Error and Root Mean Square Error as the ways to measure our prediction.

Mean Absolute Error (MAE)
n MAE -MAE = ΣY; -; | n i=1 i
(1)
Root Mean Square Error (RMSE)

R M S E = 1 ( )2
(2) RMSE=
N 1 i=1 N
(
y

i
)
2
(Where:
y i
= actual observed gas concentration

= predicted gas
concentration N = total
number of samples

B. Numerical Results

The proposed stacked LSTM model got good results. It had an error of 0.42 ppm and a root mean square error of
0.58 ppm. This means the stacked LSTM model is very good at making predictions.

If we look at a single-layer LSTM model it did not do well. The single-layer LSTM model had an error of 0.71 ppm. Traditional threshold-based monitoring did worse with much bigger differences from what actually happened. The stacked LSTM model is clearly better than the single-layer LSTM model and traditional threshold-based monitoring, at making predictions.

Table I shows how the proposed approach does better than models when we look at the performance metrics. The proposed approach is really good at getting the answers. We can see this when we compare the proposed approach to models, in Table I.

C. Graph Analysis

The graphs that show what we think will happen really follow what the sensors are actually reading. This means the proposed model is good at catching the changes in gas levels over time. Because the model can predict what will happen a little sooner we can find out about trends early on. This is very important for sending out warnings before something bad happens with the gas concentration. The proposed model is really good, at predicting what the sensor readings will be.

D. Comparison with Existing Literature

The air quality monitoring system we are talking about does a job than other Internet of Things based systems that are already out there. It makes mistakes when it predicts the air quality and it is more reliable when it sends out alerts. This system is different from systems that only react to bad air quality after it has already happened. The special computer model we use called a stacked LSTM model can actually predict when the air quality will be bad before it happens which is really good. This is similar to what people're doing now with smart ways to monitor the environment. The air quality monitoring system is really good, at sending out warnings, which is what we need to stay safe.

E. System Efficiency

The system demonstrated low processing latency (average response time ¡ 2 seconds) and stable performance under increased sensor loads, confirming its scalability and suitability for deployment in resource-constrained rural environments.

Among all of the evaluated models, Random Forest deliv- ered the best performance with an accuracy of 0.90, along with the highest precision (0.92), recall (0.90), and F1-score (0.91). Naive Bayes showed moderate performance with an accuracy of 0.81, while SVM and ANN performed poorly with the low precision and F1-scores. Overall, Random Forest demonstrates the strongest classification capability for ASD detection.

2. Comparative Discussion:

This part looks at how the new Machine Learning-based hazardous gas monitoring platform compares to air quality monitoring methods that people have written about. We compare these methods by looking at how they can monitor the air how accurate their predictions are, how quickly they respond to changes and whether they can be used in rural and remote areas. The Machine Learning-based hazardous gas monitoring platform is being compared to methods to see which one works best.

Traditional air quality monitoring systems mostly use fixed stations and alerts that go off when something goes wrong. These systems work well in cities but they are really expensive and need a lot of equipment. They are also not very good at handling a lot of data. The problem is that they only tell us something is wrong after it has already happened. For example they only send out alerts when the bad air is already, at a level.

The new platform is different. It uses cheap Internet of

International Journal of Engineering Research and Science & Technology

Things sensors and special computer programs that can predict what might happen. This means we can watch the air quality all the time and find out about problems before they get bad. We can use air quality monitoring systems and the new platform to keep an eye on the air quality and air quality monitoring systems will be helpful. The new platform and air quality monitoring systems are related to air quality monitoring.

Internet of Things air quality monitoring systems are around. They mostly show what is happening now. They do not try to guess what will happen later. These systems are good for knowing what is going on at the moment. They do not help us get ready for bad things that might happen in the future. The new platform we are talking about does more than just show what is happening now. It uses a kind of computer model called Long Short-Term Memory to predict what the air will be, like later. This means we can warn people earlier and respond faster when Internet of Things air quality is bad. Internet of Things air quality monitoring is very important. The new platform makes it better.

The stacked LSTM architecture used in this work does a job than traditional single-layer LSTM or statistical forecasting models. It is really good at picking up patterns that happen over time. When we look at the results of our experiments we see that the stacked LSTM architecture has mistakes in its predictions and sends out fewer false alarms. This shows that using deep time-series learning is a way to get accurate results. The stacked LSTM architecture is particularly good, at this because it can learn from the data in a detailed way.

Overall, the proposed system offers a balanced solution in terms of accuracy, scalability, and cost-effectiveness, making it particularly suitable for deployment in rural and remote areas where reliable environmental monitoring infrastructure is limited.

3. Visualization and Output:

Fig. 1. User Login

Fig. 2. Hazardous Gas Analysis Result

Fig. 3. Hazardous Gas Analysis Result

VII. CONCLUSION

This paper presented a Machine Learning-based digital platform for the real-time monitoring and prediction of hazardous gases in rural and remote areas. The proposed system integrates Internet of Things (IoT) technology with advanced time-series prediction models to address the limitations of conventional air quality monitoring approaches, which are often costly, reactive, and unsuitable for geographically isolated regions.

By employing a stacked Long Short-Term Memory (LSTM) architecture, the platform effectively captures temporal patterns in gas concentration data and accurately predicts future hazardous conditions. Experimental results demonstrate that the proposed system achieves high prediction accuracy, low error rates, and reliable alert generation, enabling early warnings and timely interventions. This predictive capability significantly enhances public health protection by reducing exposure risks and supporting proactive decision-making.

The platform's scalable and low-cost design makes it well suited for large-scale deployment in rural and underserved communities where continuous monitoring infrastructure is limited. Overall, the proposed solution contributes to improved environmental safety, enhanced healthcare response, and better risk management.

Future work will focus on integrating additional environmental parameters, expanding deployment across diverse geographic regions, and exploring advanced deep learning and edge computing techniques to further improve system efficiency and real-time performance.

Fig. 4. Model Prediction

REFERENCES

b1[1] [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.

[1] [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 Proceedings of the 2022 4th Asia Pacific Information Technology Conference, pp. 208–214, 2022, doi: 10.1145/3512353.3512384.

[2] [3] B. Silva, J. Rodrigues, A. Ramos, K. Saleem, I. Torre, and R. Rabeˆlo, "A mobile health system to empower healthcare services in remote regions," in Proceedings of the IEEE International Conference on E-Health Networking, Application Services (HealthCom), pp. 1–6, 2019.

[3] [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

[4] [5] World Health Organization, WHO Handbook on Indoor Radon: A Public Health Perspective, Geneva, Switzerland: World Health Organization, 2009.

[5] Nandigama, N. C. (2018). Deep Vision Networks for Multimodal Biometric Authentication: A Hybrid Feature-Level Fusion Approach with Machine Learning Optimization. Research Journal of Nanoscience and Engineering, 2(4), 22–29. https://doi.org/10.22259/2637-5591.0204005.

[6] Rongali, L. P. (2025). The Trinity of Cybersecurity: DevSecOps, Cloud Security and Cryptography in The Digital Age. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.5229585

[7] Lakshmi Prasad Rongali. (2025). Integrating AI and Devops Practices to Develop Cybersecurity Frameworks That Enhance Resilience in Utility Infrastructure. Journal of Informatics Education and Research, 5(2). https://doi.org/10.52783/jier.v5i2.2838

[8] Todupunuri, A. (2024). Develop Machine Learning Models to Predict Customer Lifetime Value for Banking Customers, Helping Banks Optimize Services. International Journal of All Research Education & Scientific Methods, 12(10), 1254–1259. https://doi.org/10.56025/ijaresm.2024.1210241254

[9] Todupunuri, A. (2024). Explore How AI Can Be Used To Create Dynamic And Adaptive Fraud & Rules That Improve The Detection And Prevention Of Fraudulent & Activities In Digital Banking. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.5014699

[10] Nandigama, N. C. (2019). Hybrid Facial Recognition System Using Histogram of Oriented Gradients and Deep Learning with Dimensionality Reduction. Research Journal of Nanoscience and Engineering, 3(4), 30–35. https://doi.org/10.22259/2637-5591.0304005

[11] Nandigama, N. C. (2016). Scalable Suspicious Activity Detection Using Teradata Parallel Analytics And Tableau Visual Exploration