# Voting Ensemble Learning Model (VELM) for Harmful Gas Detection in Environmental Applications

**Oluwabukola F. Ajayi[1], Alfred A. Udosen[2*], Wumi Ajayi[3], Blaise O. Ohwo[4] and Afolarin I. Amusa[5]**

[1,2,4&5]Department of Computer Science, [3]Department of Software Engineering,
School of Computing, Babcock University, Ilishan-Remo, Nigeria
E-mail: otusileo@babcock.edu.ng, ajayiw@babcock.edu.ng, ohwoo@babcock.edu.ng, amusaa@babcock.edu.ng
*Corresponding Author: udosena@babcock.edu.ng

*Abstract* - The rapid industrialization and technological advancements of recent decades have significantly improved various facets of human life but have also intensified air pollution through the emission of harmful gases. Accurate detection of these pollutants, including Carbon Monoxide (CO), Ammonia ($NH_3$), and Sulfur Dioxide ($SO_2$), is critical for mitigating environmental and health hazards. Traditional sensor technologies often fail due to calibration issues and susceptibility to extreme temperatures, underscoring the need for advanced monitoring solutions. This study investigates the integration of artificial intelligence (AI), particularly ensemble machine learning models, to address these challenges. Leveraging the Gas Sensor Array Drift Dataset, a Voting Ensemble Learning Model (VELM) was developed and benchmarked against individual machine learning classifiers such as Random Forest, Support Vector Machine, and Logistic Regression. The VELM demonstrated superior performance, achieving a classification accuracy of 99.46%, surpassing conventional methods while maintaining low variance. Despite marginal accuracy differences with Random Forest, VELM's majority voting approach consistently ensured robust performance. The findings highlight the transformative potential of ensemble learning in environmental monitoring and provide a foundation for future research. Recommendations include exploring deep learning enhancements and deploying the model in real-world settings to refine its applicability for detecting a broader range of harmful gases, thereby advancing public safety and environmental sustainability.

*Keywords:* Air Pollution, Ensemble Machine Learning, Voting Ensemble Learning Model (VELM), Gas Sensor Array Drift Dataset, Environmental Monitoring

## I. INTRODUCTION

The recent global increase in industrialization and technological advancement activities has led to improvements in all spheres of life. The burning of fossil fuels, such as coal, oil, and gas, for heating and transportation contributes to air pollution, global warming, climate change, and acid rain [1]. Gas sensors are developed using mathematical equations embedded within microcontrollers to compute gas concentrations, enabling timely warnings and interventions [2].

The process involves collecting gas level readings with sensors, such as MQ2 to MQ135, transmitting data to microcontrollers, and storing it in databases. Additionally, thermal cameras capture infrared radiation images for comprehensive analysis, facilitating the detection of harmful gases in the air [3]. Some of these harmful gases, as noted in [4], include Carbon Monoxide (CO), Hydrogen ($H_2$), Ammonia ($NH_3$), Hydrogen Sulfide ($H_2S$), Acetone ($CH_3COCH_3$), Ethanol ($C_2H_5OH$), Sulfur Dioxide ($SO_2$), and Nitrogen Dioxide ($NO_2$).

However, the need to effectively monitor harmful gases released into the air due to industrial activities cannot be overemphasized. This is attributed to the limitations of sensor technology, such as failure in detecting harmful gases due to high temperatures and poor calibration, among other factors [5]. Owing to the recorded success of artificial intelligence, N. N. Viet *et al.,* [4] conducted a study aimed at developing efficient AI tools for gas detection and categorization, focusing on designing a simple sensor array configuration and comparing machine learning models. A gradient-boosting model, which achieved 100% accuracy on training datasets, demonstrated its potential for gas-sensing applications.

The exponential increase in industrialization necessitates a more robust approach to detecting harmful gases. According to A. A. Udosen *et al.,* [6], the notable successes recorded with voting ensemble learning models, an artificial intelligence (AI) algorithm, highlight their potential for exploitation. A. A. Abiona *et al.,* [7] and U. A. Umoh *et al.,* [8] argued that AI improves decision-making and optimizes output. Additionally, T. Deep Singh *et al.,* [9] stated that machine learning enables computers to learn automatically and enhances their performance without explicit programming. Despite its wide acceptance, E. Onuiri *et al.,* [10] emphasized that regulations should be placed on the use of machine learning and related AI algorithms.

## II. LITERATURE REVIEW

Artificial Intelligence (AI), as defined by U. J. Nzenwata *et al.,* [11], is a scientific effort to develop devices capable of replicating human intellect. Machine learning, a subset of AI, is a mathematically constructed technique enabling

intelligent learning using refined datasets. As part of the tasks AI can automate, A. Ahmadi [12] highlights its capacity for logic, interpretation of meaning, drawing broad conclusions, and learning from experience. Leveraging these capabilities, B. Vaferi et al., [1] developed a stacked machine learning model that used zinc oxide-based sensors to detect hydrogen. The authors utilized XGBoost and Extra Trees regressors to create a highly accurate and generalizable approach, outperforming traditional artificial neural networks in performance metrics.

Furthermore, N. N. Viet et al., [4] compared the performance of several machine learning models in identifying harmful gases through thermal fingerprint measurements and Metal Oxide Semiconductor (MOS) sensors. Additionally, S. Mu et al., [13] implemented an electronic nose system combining a MOS-based micro-electromechanical system (MEMS) gas sensor array and a one-dimensional convolutional neural network (1D-CNN) to identify several harmful gases. Unlike conventional machine learning algorithms, which achieved an accuracy

of approximately 80%, the developed model performed with about 99.8% accuracy. However, the model was limited to detecting only seven gases.

To enhance kitchen safety, K. Kumar et al., [14] developed an Internet of Things (IoT) and Machine Learning (ML) integrated Hazardous Gas Detection System (IoT-HGDS) using ESP32 MOX-based sensors and DHT22. The system was designed to mitigate the risk of fire hazards in the kitchen, thereby reducing the likelihood of severe damage or injury. According to E. E. Onuiri et al., [15], the success of these models is primarily based on the precision capabilities of machine learning algorithms.

To produce a diverse dataset of harmful gases, J. Fonollosa et al., [5] conducted an experiment that generated 13,910 records of six harmful gases from chemical sensors at different concentration levels. Figure 1 illustrates the experimental setup used for data acquisition via a computer-supervised continuous flow system.
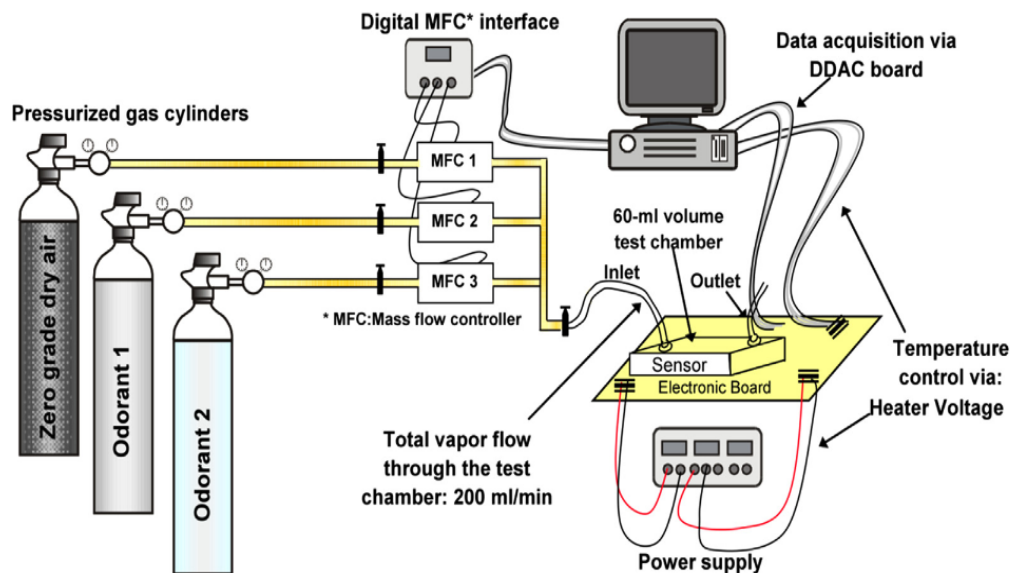


Fig. 1 Experimental setup used for data acquisition

### III. METHODOLOGY

*A. Dataset*

The study utilized the Gas Sensor Array Drift Dataset obtained from the Data Hub repository [16]. This dataset contains 13,910 rows of data with 128 features and an output label representing six classes of harmful gases, derived from the concentration levels of an array sensor [17]. The data was collected over approximately 36 months (January 2007 to February 2011) using $SnO_2$ Taguchi Gas Sensors (TGS) at the BioCircuits Institute's ChemoSignals Laboratory, University of California, San Diego.

The dataset was recorded using 16 gas sensors arranged in the following matrix order: TGS2602, TGS2602, TGS2600,

TGS2600, TGS2610, TGS2610, TGS2620, TGS2620, TGS2602, TGS2602, TGS2600, TGS2600, TGS2610, TGS2610, TGS2620, TGS2620. However, the dataset exhibited no clear pattern, making it challenging for humans to interpret.

This harmful gas detection model integrates five classification algorithms to identify gases such as Ethanol, Ethylene, Ammonia, Acetaldehyde, Acetone, and Toluene. The model progresses through stages including data cleaning, feature selection, and ensemble learning. Five models are combined to create a heterogeneous ensemble model, with a weighted average used for prediction.

The class labels were used to encode the detected harmful gases in the dataset, which was saved in the comma-

separated value (CSV) format. The dataset comprises 13,910 samples with 128 features and a multi-classification label, as encoded in Table I.

TABLE I MULTI-CLASSIFICATION LABEL

| Gas Detected | Label | Freq |
|---|---|---|
| Ethanol | 1 | 2565 |
| Ethylene | 2 | 2926 |
| Ammonia | 3 | 1641 |
| Acetaldehyde | 4 | 1936 |
| Acetone | 5 | 3009 |
| Toluene | 6 | 1833 |

*A. Preparation Stage*

Before model development, the dataset undergoes preparation, involving two key stages:

1. *Pre-Processing:* Data cleaning addresses missing values and outliers caused by poor sensor readings. Standardization transforms the dataset to have attributes with a Gaussian distribution, facilitating the application of modeling techniques. The StandardScaler class from Scikit-learn in Python is commonly used for this purpose.
2. *Feature Selection:* Identifying essential features enhances model performance by excluding irrelevant ones. With 128 independent features in the dataset, Principal Component Analysis (PCA) is employed to select only the necessary features, optimizing model performance.

*B. Training Stage*

During the training stage, the dataset is utilized to train models using K-Nearest Neighbors (KNN), Decision Tree (DT), Logistic Regression (LR), Support Vector Machine (SVM), and Random Forest (RF) classification techniques. Subsequently, a Voting Ensemble Learning Model (VELM) is developed to enhance efficiency in dynamic environments. The dataset is split into 80% for training and 20% for testing and validation. Assessing the accuracy of each model during training provides insight into their performance.

*C. Validation and Testing Stage*

In the validation and testing stage, the remaining 20% of the dataset is used to evaluate model performance. K-fold cross-validation, with kkk values of 5 or 10, is employed to split the dataset for training and testing, ensuring more reliable performance estimates.

Evaluation metrics, including classification accuracy, confusion matrix, and classification report, are utilized to accurately gauge model performance on fresh data.

*D. Model Evaluation*

In the model evaluation phase, the effectiveness of the developed harmful gas detection model is assessed using various machine learning classification performance metrics. These include:

1. *Classification Accuracy:* Measures the proportion of correctly predicted events.

2. *Confusion Matrix:* Provides insights into the model's accuracy across different classes.

3. *Classification Report:* Offers a comprehensive analysis of classification performance, including precision, recall, and F1-score.

Evaluation involves spot-checking six machine learning algorithms using 10-fold cross-validation, with mean accuracy measures indicating algorithm performance. Ethical considerations ensure adherence to standard rules and guidelines. The expected contribution to knowledge lies in the utilization of the Voting Ensemble Learning Model (VELM) for detecting harmful gases. Table II shows the confusion matrix for the data points.

TABLE II CONFUSION MATRIX

| | Predicted No: 0 | Predicted Yes:1 |
|---|---|---|
| Actual No: 0 | TP | TN |
| Actual Yes: 1 | FP | FN |

## IV. RESULTS AND DISCUSSION OF FINDINGS

Ensemble learning techniques, including the Voting Ensemble Learning Model (VELM), were applied for harmful gas detection, demonstrating significant promise in environmental science and engineering. The study compared VELM with robust base classifiers such as Support Vector Machine (SVM), Random Forest (RF), Logistic Regression (LR), Decision Trees (DT), and K-Nearest Neighbors (KNN). Data preparation and model training were conducted, with metrics such as recall and F1-score calculated for performance evaluation. Results indicated that VELM outperformed individual models, highlighting its superior efficacy compared to existing models.

*A. Data Preparation*

The Gas-Drift dataset from the DataHub repository was utilized, consisting of 13,910 records gathered from an array of 16 metal-oxide gas sensors, with 128 distinct features. Data preparation involved a scaling transformation to address negative values arising from sensor malfunctions. Scaling ensured uniformity across all values, standardizing them with a mean of 0 and a standard deviation of 1.

Oluwabukola F. Ajayi, Alfred A. Udosen, Wumi Ajayi, Blaise O. Ohwo and Afolarin I. Amusa

Following data transformation, each attribute's values achieved a mean of 0 and a standard deviation of 1 [28]. Feature selection was conducted using Principal Component Analysis (PCA) to identify the number of principal components, reducing the dataset to 20 principal components.

*B. Model Training and Testing*

In supervised learning, the train-test split was used to evaluate model performance on new data, with x representing 128 independent variables and y representing multi-class target values.

The dataset underwent K-fold cross-validation, splitting it into 80% for training and 20% for testing subsets. Each base classifier contributed to the Voting Ensemble Learning Model, with the weighted average used to determine the model's accuracy score.

*C. Evaluation of Model Performance Metrics*

The study evaluated model performance using three primary criteria:

*D. Classification Accuracy*

Classification accuracy was used to quantify the proportion of correct predictions made out of all predictions. Multiplying the value by 100 provided the accuracy in percentage terms. Table III illustrates the accuracy of the models following evaluation.

TABLE III CLASSIFICATION ACCURACY OF EACH MODEL

| Sl. No. | Model | Accuracy (%) | Standard Deviation (%) |
|---------|-------|--------------|------------------------|
| 1 | VELM | 99.46 | 0.1774 |
| 2 | RF | 99.44 | 0.2199 |
| 3 | KNN | 99.24 | 0.208 |
| 4 | LR | 99.02 | 0.2612 |
| 5 | SVM | 98.32 | 0.4490 |
| 6 | DT | 97.33 | 0.3291 |

The results from the study showed that the developed heterogeneous Voting Ensemble Learning Model (VELM) demonstrated a marginal difference in accuracy compared to the homogeneous Random Forest (RF) machine learning model. The VELM outperformed the other models, with Decision Tree (DT) being the poorest-performing model. The classification accuracies for VELM, RF, and DT were 99.46%, 99.44%, and 97.33%, respectively.

From the classification accuracy results, it can be concluded that the VELM was the best-performing model. Higher accuracy levels of the evaluated models indicate better performance in detecting harmful gases. Additionally, as the standard deviation from the mean decreased for each evaluated model, the accuracy performance of the models improved.

Figure 2 illustrates the Boxplot comparison of conventional machine learning techniques, with RF having the best prediction estimate.
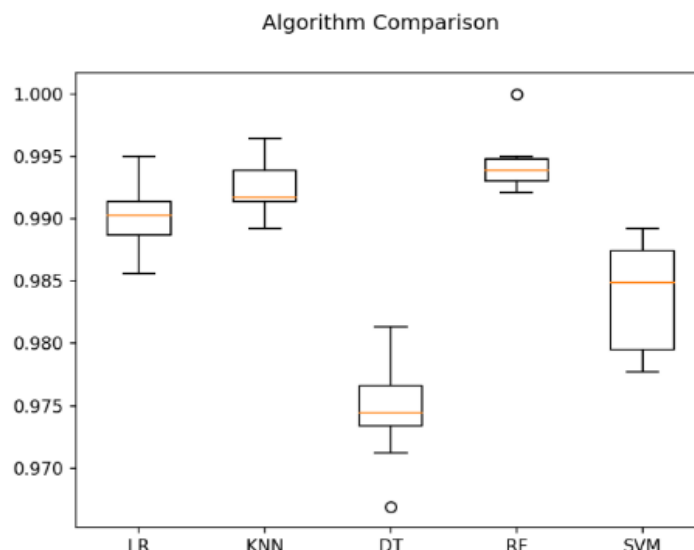


Fig. 2 Boxplot model comparison

*E. Confusion Matrix*

The confusion matrix was used to compare the accuracy of the multi-class models.

TABLE IV CONFUSION MATRIX SHOWING DATA POINTS IN THE TRUE POSITIVE REGION FOR CLASS 1

| Sl. No. | Model | Data Points in the True Positive region |
|---------|-------|------------------------------------------|
| 1 | VELM | 522 |
| 2 | RF | 521 |
| 3 | KNN | 520 |
| 4 | LR | 517 |
| 5 | SVM | 515 |
| 6 | DT | 512 |

The results from the study, as shown in Table IV, indicated that the developed Voting Ensemble Learning Model (VELM) had the highest number of values in the True Positive category, followed by Random Forest (RF), with Decision Tree (DT) having the lowest.

Here, data points in the True Positive region for Class 1 tells how each of the techniques performed.

*F. Performance Analysis*

The Voting Ensemble Learning Model (VELM) was benchmarked against classic machine learning classifiers, including Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Decision Tree (DT), and Logistic Regression (LR), to determine optimal training/testing scenarios. The results are displayed in Figure 3.

The classification report indicated that the VELM achieved the highest accuracy, with a weighted average classification accuracy of 99% and an average F1-score of 1%.
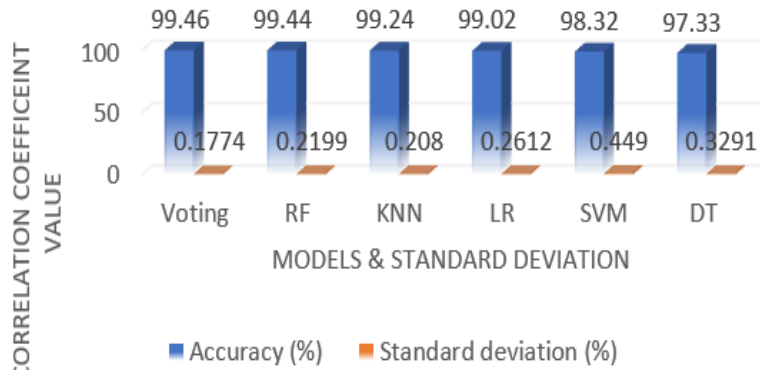


Fig. 3 Models' Performance Estimation

In the study both Random Forest and the developed Voting Ensemble Learning Model (VELM) outperformed individual single-based classic machine learning models. The VELM achieved the highest training accuracy of 99.46% with a low standard deviation of 0.1774%. By leveraging the weighted average of different base classifiers, the VELM addressed weaknesses to ensure superior performance. Although there was a marginal 0.02% difference compared to Random Forest, VELM's majority voting approach improved overall performance. Table V compares the performance of the developed VELM with previously developed ensemble learning models, demonstrating the superior performance of the VELM over other models.

TABLE V BENCHMARKING THE ENSEMBLE LEARNING MODEL

| Sl. No. | Method | ACC | Precision | Sensitivity | Specificity | F1-Score |
|---------|--------|-----|-----------|-------------|-------------|----------|
| 1 | VELM | 99.46% | 99.01% | 99.01% | 99.01% | 99.01% |
| 2 | [12] | 97.06% | 98.73% | 97.17% | 99.24% | 97.94% |
| 3 | [46] | 92.49% | 92.20% | 94.93% | 91.13% | 93.55% |
| 4 | [52] | 88.63% | 90.28% | 86.09% | 94.35% | 88.14% |

## V. CONCLUSION

This study addresses the critical need for accurate harmful gas detection using a Voting Ensemble Learning Model (VELM) integrating IoT technology and machine learning algorithms. By combining five well-performing classifiers, including Support Vector Machine (SVM), Logistic Regression (LR), Decision Trees (DT), K-Nearest Neighbors (KNN), and Random Forest (RF), the VELM achieves high prediction accuracy. Results indicate that VELM and Random Forest outperform other individual algorithms, with VELM demonstrating marginal superiority with a 99.46% accuracy score and a standard deviation of 0.1774%. Benchmark analysis confirms VELM's superior

performance, suggesting its potential for enhancing gas detection accuracy. Future research could explore advanced techniques such as ensemble learning and deep learning with larger, high-quality datasets. The study utilized the Gas Sensor Array Drift Dataset, containing 13,910 rows and 128 features, and introduced a Voting Ensemble Learning Model (VELM) that combines SVM, Logistic Regression, Decision Trees, KNN, and Random Forest algorithms. The VELM achieved an impressive accuracy of 99.46%, making it suitable for gas detection systems. This underscores the need for advanced machine learning techniques, such as ensemble learning, to overcome the limitations of classic methods. The VELM offers a promising solution for real-time harmful gas detection, addressing critical safety concerns. The methodology employed in this study for detecting harmful gases could be applied to various detection systems by prioritizing essential features. Pilot studies involving local stakeholders could enhance data collection, particularly in regions with data scarcity, aiding in system improvement and implementation.

*A. Contribution to Knowledge*

The study addresses the limitations of classic machine learning algorithms in gas detection by proposing a Voting Ensemble Learning Model (VELM) to overcome issues such as overfitting and local minima. The developed VELM offers the potential for real-time detection of harmful gases, thereby potentially saving lives. Additionally, it opens avenues for further academic research to enhance its outcomes.

*B. Suggestions for Further Studies*

Future research could explore the implementation and testing of enhanced ensemble models integrating deep learning algorithms for detecting harmful gases.

*C. Limitation of the Study*

Accessing appropriate datasets and computational tools posed challenges, slowing down model development for harmful gas detection.

## REFERENCES

[1] B. Vaferi, M. Dehbashi, A. Khandakar, M. A. Ayari, and S. Amini, "Development of a stacked machine learning model to compute the capability of ZnO-based sensors for hydrogen detection," *Sustainable Materials and Technologies*, vol. 39, p. e00863, Apr. 2024, doi: 10.1016/j.susmat.2024.e00863.

[2] W. L. Hsu, C. Y. Ho, C. K. Liang, Y. C. Shiau, H. N. Hsieh, and S. C. Lai, "Application of IoT in the Prevention of Carbon Monoxide Poisoning," *Sensors and Materials*, vol. 31, no. 11, p. 3465, Nov. 2019, doi: 10.18494/SAM.2019.2482.

[3] A. Adekunle, I. I. Umanah, K. E. Ibe, and I. M. Rukewe, "Statistical Analysis of Fire Outbreaks in Homes and Public Buildings in Nigeria: A Case Study of Lagos State," *International Journal of Engineering Research and Advanced Technology*, vol. 4, no. 8, pp. 21-30, 2018, doi: 10.31695/IJERAT.2018.3294.

[4] N. N. Viet, P. H. Phuoc, L. V. Thong, N. V. Chien, and N. Van Hieu, "A comparative study of machine learning models for identifying noxious gases through thermal fingerprint measurements and MOS sensors," *Sensors and Actuators A: Physical*, vol. 375, p. 115510, Sep. 2024, doi: 10.1016/j.sna.2024.115510.

[5] J. Fonollosa, I. Rodríguez-Luján, and R. Huerta, "Chemical gas sensor array dataset," *Data in Brief*, vol. 3, pp. 85-89, Jun. 2015, doi: 10.1016/j.dib.2015.01.003.

[6] A. A. Udosen, O. F. Ajayi, and J. O. Adelowo, "Voting Ensemble Learning Model (VELM) for Harmful Gas Detection Systems: A Proposed Model," *Current Trends in Information and Communication Technology Research*, vol. 2, no. 1, pp. 1-12, 2023.

[7] A. A. Abiona, M. O. Olarinde, and M. O. Ajinaja, "Enhanced Donkey and Smuggler Optimization Algorithm for Holistic Student Admissions in Polytechnics," *Asian Journal of Computer Science and Technology*, vol. 13, no. 2, pp. 56-60, Nov. 2024, doi: 10.70112/ajcst-2024.13.2.4302.

[8] U. A. Umoh and A. A. Udosen, "Sugeno-Type Fuzzy Inference Model for Stock Price Prediction," *International Journal of Computer Applications*, vol. 103, no. 3, 2014, doi: 10.5120/18051-8957.

[9] T. Deep Singh and R. Bharti, "Detection and Classification of Plant Diseases in Crops (*Solanum lycopersicum*) due to Pests Using Deep Learning Techniques: A Review," *Asian Journal of Computer Science and Technology*, vol. 12, no. 2, pp. 39-47, Nov. 2023, doi: 10.51983/ajcst-2023.12.2.3735.

[10] E. Onuiri, B. G. Akwaronwu, and K. C. Umeaka, "Environmental and Genetic Interaction Models for Predicting Lung Cancer Risk Using Machine Learning: A Systematic Review and Meta-Analysis," *Asian Journal of Computer Science and Technology*, vol. 13, no. 1, pp. 45-58, Apr. 2024, doi: 10.70112/ajcst-2024.13.1.4266.

[11] U.-J. Nzenwata, A. G. Abiodun, A. Olayinka, O. J. Adeniyi, and A. B. Gazie, "Parkinson's Disease Prediction Using Convolutional Neural Networks and Hand-Drawn Image Analysis," *Asian Journal of Computer Science and Technology*, vol. 13, no. 2, pp. 1-13, Sep. 2024, doi: 10.70112/ajcst-2024.13.2.4270.

[12] A. Ahmadi, "Unravelling the Mysteries of Hallucination in Large Language Models: Strategies for Precision in Artificial Intelligence Language Generation," *Asian Journal of Computer Science and Technology*, vol. 13, no. 1, pp. 1-10, Mar. 2024, doi: 10.70112/ajcst-2024.13.1.4144.

[13] S. Mu, W. Shen, D. Lv, W. Song, and R. Tan, "Inkjet-printed MOS-based MEMS sensor array combined with one-dimensional convolutional neural network algorithm for identifying indoor harmful gases," *Sensors and Actuators A: Physical*, vol. 369, p. 115210, Apr. 2024, doi: 10.1016/j.sna.2024.115210.

[14] K. Kumar, A. Verma, and P. Verma, "IoT-HGDS: Internet of Things integrated machine learning based hazardous gases detection system for smart kitchen," *Internet of Things*, vol. 28, p. 101396, Dec. 2024, doi: 10.1016/j.iot.2024.101396.

[15] E. E. Onuiri, C. Ogbonna, and K. C. Umeaka, "Performance of Predictive Models in Cervical Cancer Recurrence: A Systematic Review and Meta-Analysis of Biomarkers and Prognosis," *Asian Journal of Computer Science and Technology*, vol. 13, no. 2, pp. 14-28, Oct. 2024, doi: 10.70112/ajcst-2024.13.2.4271.

[16] UC Irvine Machine Learning Repository, "Gas Sensor Array Drift at Different Concentrations." Available: https://datahub.io/machine-learning/gas-drift.

[17] W. Elmenreich, "A review on system architectures for sensor fusion applications," in *IFIP International Workshop on Software Technologies for Embedded and Ubiquitous Systems*, Springer, 2007, pp. 547-559.